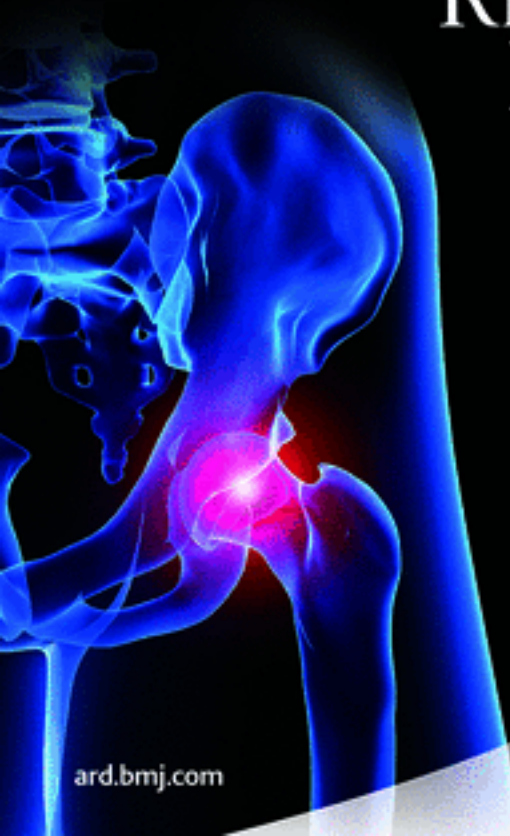
May 2017 Volume 76 Issue 5

Impact Factor 12.384

Annals of the Rheumatic Diseases

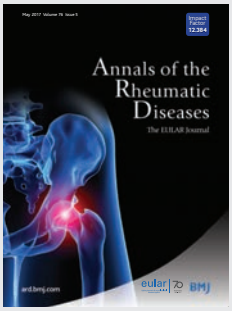
The EULAR Journal



ard.bmj.com

eular 70 BMJ

ARD
The Eular Journal



Editorial office
Annals of the Rheumatic Diseases
BMJ Publishing Group Ltd
BMA House
Tavistock Square
London WC1H 9JR,UK
T: **+44 (0)20 7383 6250**
F: **+44 (0)20 7383 6668**
E: **ard@bmj.com**
Twitter: **@ARD_BMJ**
ISSN: 0003-4967 (print)
ISSN: 1468-2060 (online)
Impact Factor: 12.384

ARD is published by BMJ Publishing Group Ltd typeset by Exeter Premedia Services Private Ltd, Chennai, India and printed in the UK on acid-free paper.

Annals of the Rheumatic Diseases (ISSN No: 0003-4967) is published monthly by BMJ Publishing Group and distributed in the USA by Air Business Ltd. Periodicals postage paid at Jamaica NY 11431 POSTMASTER: send address changes to *Annals of the Rheumatic Diseases*, Air Business Ltd, c/o Worldnet Shipping Inc., 156-15, 146th Avenue, 2nd Floor, Jamaica, NY 11434, USA.

Editorial

Criteria

Recommendation

Clinical and epidemiological research

MORE CONTENTS ▶

Basic and translational research

Electronic pages

Correspondence

Lesinurad combination therapy with allopurinol in gout: do CLEAR studies make the treatment of gout clearer?

Jasvinder A Singh[1,2,3]

Gout is an often forgotten disease,[1] despite being the most common inflammatory arthritis in adults in the Western world.[2–4] Lesinurad, a new urate-lowering therapy (ULT), is now approved for the treatment of gout in the USA and the European Union,[5 6] and is in phase III programmes as a combination therapy in many other countries. Lesinurad is a selective inhibitor of urate/anion exchanger 1 and organic acid transporter 4, two urate transporters responsible for the reabsorption of urate from the proximal renal tubule.[7] Probenecid and benzbromarone are the other uricosurics available for use in gout as monotherapy or combination with allopurinol in case of an inadequate response to allopurinol. In the current issue of the journal, Bardin *et al* present the results of a 12-month, randomised, phase III trial efficacy and safety of oral lesinurad (200 or 400 mg) in combination with allopurinol.[8] Combining Lesinurad with Allopurinol in Inadequate Responders (CLEAR) were two replicate studies, one in the USA (CLEAR-1)[9] and one in Europe (CLEAR-2), published in the current issue of this journal.[8] CLEAR-2 compared daily lesinurad (200 or 400 mg orally) with placebo when added to allopurinol in 610 patients with gout with serum uric acid (sUA) above target (<6 mg/dL or <0.36 mmol/L) and frequent gout flares (≥2 gout flares in the prior year).

WHAT DID THE CLEAR-2 TRIAL RESULTS SHOW?

The primary trial endpoint of sUA <6 mg/dL (ie, <0.36 mmol/L) at 6 months was achieved by a significantly greater proportion of patients, 55% and 66% in the lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups versus 23% in the allopurinol alone group (p<0.0001 vs either lesinurad group).[8] These rates were similar to the achievement of target sUA <6 mg/dL (ie, <0.36 mmol/L) in CLEAR-1.[9] To my knowledge, CLEAR-2[8] and the replicate study, CLEAR-1,[9] are the two largest randomised control trials (RCTs) of lesinurad versus placebo in patients with symptomatic gout with frequent flares despite treatment with allopurinol.

Harms were similar in the lesinurad 200 mg versus placebo groups, but somewhat higher in the lesinurad 400 mg group. Differences were noted in the safety profile. Serious adverse events and renal adverse events occurred in similar proportions of patients receiving the lesinurad 200 mg+allopurinol and placebo +allopurinol group (4–6%), but in 2–3 times as many people in the lesinurad 400 mg+allopurinol group, with incidences of 10% and 15%, respectively. Among renal adverse events, increased serum creatinine was seen in same proportions of the lesinurad 200 mg+allopurinol and allopurinol-only groups, but in twice as many people in the higher lesinurad group. Renal failure and serum creatinine elevations of ≥1.5 times baseline value were more frequent with both lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups compared with allopurinol alone. These serum creatinine elevations were mostly reversible, except in seven patients in the lesinurad 400 mg dose and three patients in the allopurinol group; similarly, five cases of serum creatinine elevations of ≥2 times baseline value were unresolved in the lesinurad 400 mg dose versus none in the other groups by the last study visit.

Evidence of the two studies, CLEAR-1 and CLEAR-2, indicates that compared with placebo plus allopurinol, lesinurad 200 mg/day in combination with allopurinol median dose of 300 mg/day was efficacious and was not associated with a significant increase in rate of serious adverse events or renal adverse events. Lesinurad 400 mg/day in combination with allopurinol is effective, but was associated with clinically meaningfully higher rates of serious adverse events and renal adverse events and possibly renal failure.

WHAT CONCERNS REMAIN?

Allopurinol was used in 'standard-of-care' doses in patients with gout who were randomised in CLEAR-1, with a mean allopurinol dose of approximately 300 mg/day, similar to that reported in observational cohorts.[10 11] This allopurinol dose achieves target sUA <6 mg/dL in only 50% of the patients,[10 11] meaning that the standard allopurinol dose is suboptimal in half of the patients with gout. Allopurinol dose had not been titrated to a maximum approved dose of 800 mg daily (or 900 mg approved maximum dose in Europe) to achieve target sUA in CLEAR studies. The CLEAR-1 and CLEAR-2 studies show that in patients with gout who have failed a suboptimal 100–300 mg daily dose of allopurinol, lesinurad is more beneficial than placebo in achieving target sUA. These studies provide us with the first glimpse of benefits and harms of lesinurad compared with placebo. Perhaps a more informative study for clinicians would have been the use of new therapies (such as lesinurad) in patients who had truly been refractory to allopurinol, that is, failed to achieve a target sUA <6 mg/dL despite adequate allopurinol doses titrated to 800 mg of allopurinol or more and used for an adequate duration of treatment (6–12 months). Pharmaceutical companies in the field of developing urate-lowering therapies would be better off planning and conducting trials targeting these true allopurinol refractory patient populations.[12] Rheumatologists have started using the therapeutic doses of allopurinol up to the approved maximum doses (800 mg/day in the USA and 900 mg/day in Europe) in their current practice. I hope that this practice of the use of titrated therapeutic allopurinol doses will spread to internists, cardiologists, nephrologists, podiatrists and other physician colleagues, with proper education efforts.

Serum creatinine elevations ≥2 times baseline levels resolved in all patients in the lesinurad 200 mg+allopurinol group but were unresolved in five patients in the lesinurad 400 mg+allopurinol group in CLEAR-2[8] and unresolved in two cases in the lesinurad 400 mg+allopurinol group in CLEAR-1.[9] Not surprisingly, in CLEAR-1, renal failure occurred in 1% and 1.5% of patients receiving lesinurad 200 and 400 mg versus 0.5% in placebo, respectively. This is clinically relevant. The absolute risk difference for renal failure with lesinurad 400 mg daily versus placebo was 1% and the relative risk was approximately three times. Was there anything peculiar about these patients

[1]Birmingham VA Medical Center, Birmingham, Alabama, USA; [2]Department of Medicine at the School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA; [3]Division of Epidemiology at the School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA

Correspondence to Dr Jasvinder A Singh, University of Alabama, Faculty Office Tower 805B, 510 20th Street S, Birmingham, AL 35294, USA; Jasvinder.md@gmail.com

regarding comorbidity and/or concomitant medications? Can the clinician be better guided as to who these patients might be so that they can avoid the higher dose of lesinurad in these patients or try some other strategy? It would be very informative to see detailed characteristics of these patients in subsequent publication (s) so that clinicians and patients are better informed about who might be at a higher risk of an irreversible creatinine elevation or renal failure with the higher lesinurad dose.

WAS THERE ANY EVIDENCE OF DIMINISHING CLINICAL RESPONSE OVER TIME? WHAT ELSE DOES THIS TRIAL TEACH US?

The success rates for achieving target sUA <6 mg/dL with lesinurad co-therapy at doses 200 and 400 mg/day were 55% and 65% at 6 months compared with 22% with placebo; proportions were 49% versus 55% versus 26% at 12 months, showing slight loss of efficacy in lesinurad arms. The dropout rate was non-differential between 6 and 12 months between arms and minimal, 4 versus 1 versus 5 people, respectively. Most patients dropped out before the 6-month time point at 21% and 25% versus 23%, respectively. Thus, the slight loss in efficacy in lesinurad arms at 12 months is unlikely to be explained by loss to follow-up. Whether a 6–10% difference in proportions with target sUA between 6 and 12 months is a chance finding or a reduction of lesinurad efficacy or adherence remains to be seen.

Even though the primary sUA outcome was significantly better in the lesinurad groups compared with placebo, the tophi resolution rate was similar to that observed in the placebo group—31% and 28% versus 33% had complete resolution of ≥1 target tophi. Was the study too short or patients too few to show differences in tophi resolution? Is the tophus measure not sensitive to change in this patient population? Certainly, tophi were secondary outcome, and therefore, the current study was not powered to detect significant between-group differences. Studies that recruit a larger number or proportion of gout patients with tophi are needed to demonstrate whether tophi resolution is faster in patients with lesinurad compared with placebo. I suspect that such study will be of longer duration since tophi size reduction and resolution cannot be expected in a short duration efficacy trial.

An interesting study finding was that up to a third of the patients treated with allopurinol alone had resolution of ≥1 target tophi, demonstrating that allopurinol is effective, even when used in suboptimal doses. This finding is similar to 50% reduction in tophus area with allopurinol dose of 200–300 mg daily in another RCT.[13] If allopurinol 300 mg daily dose, which is effective in achieving target sUA in 50% of the patients, leads to 29–50% reduction in target tophi, how much reduction will one expect with a 100% effective dose? Is it possible that we already have a cure for gout and just have not realised it yet? I suspect that if we use allopurinol at an optimal final dose (up to 800 mg/day in the USA; 900 mg/day in Europe) uptitrated from a lower, starting dose to achieve target sUA, it will be an extremely effective tool both for lowering sUA and resolving tophi. Of course, availability of other ULTs such as febuxostat, pegloticase, traditional uricosurics and now lesinurad, where titrated allopurinol does not succeed or is contraindicated, can help improve our success rate of curing gout.

Importantly, and quite interestingly, a very small proportion of patients (<15%) had gout flares requiring treatment since all received anti-inflammatory prophylaxis, which started before the initiation of lesinurad or placebo, and continued for 5 months. The proportion with gout flares did not differ across the lesinurad versus placebo groups. ULT initiation without an anti-inflammatory prophylaxis is associated with frequent gout flares. This study demonstrates that anti-inflammatory prophylaxis for gout flares is effective in preventing acute flares if started early and taken for a few months when an ULT is started.

The 25% discontinuation rate in this published study,[8] although similar to the 33% discontinuation rate in a pivotal RCT of febuxostat,[13] is still higher than the acceptable 20% dropout rate in trials. A high dropout rate, defined as >20%, puts the RCT results at a higher risk of bias,[14] a study limitation that must be considered while interpreting the trial results and making conclusions.

HOW WILL LESINURAD'S AVAILABILITY CHANGE THE MANAGEMENT OF GOUT?

I think this is the golden era for the treatment of gout. New therapies are being approved and becoming available for the treatment of hyperuricemia and acute gout flares. Investigations of new mechanisms for urate lowering reflect a better understanding of urate metabolism and novel pathways.

Lesinurad's availability offers a new choice of a uricosuric ULT co-therapy option to be used concomitantly with allopurinol for patients with refractory gout despite use of allopurinol. Interestingly, gout flares and tophi resolution were not significantly different from placebo in this 12-month study, but this may be related to small sample size, adequate anti-inflammatory prophylaxis or a small effect size with lesinurad. Based on the evidence, one can consider lesinurad 200 mg/day as a second-line therapeutic option to be co-administered with allopurinol after an adequately titrated dose of allopurinol (approved up to 800 mg/day in the USA and 900 mg/day in Europe) has been tried and fails to control sUA, frequent flares or tophi.

WHAT SHOULD A CLINICIAN MONITOR WHILE TREATING PATIENTS WITH LESINURAD?

The data presented in the current study and CLEAR-1 provide confidence regarding the risk/benefit ratio of lesinurad 200 mg daily dose in combination with allopurinol. However, in patients without a normal renal function, one needs to cautiously evaluate and discuss the risk/benefit ratio of lesinurad 400 mg in combination with allopurinol, given a higher risk of serious adverse events and renal adverse events, including renal failure, compared with placebo in both CLEAR-1 and CLEAR-2. It seems that regular monitoring of renal function is prudent when starting lesinurad.

WHAT IS THE TAKE HOME MESSAGE?

CLEAR-1 and CLEAR-2 studies bring new knowledge for clinicians and a new treatment for patients. First, lesinurad 200 mg in combination with allopurinol is an effective and safe option for patients with symptomatic gout despite an average allopurinol dose of 300 mg/day. Second, this 12-month study of lesinurad showed an important improvement in sUA target achievement compared with placebo, but no significant difference in gout flares or gouty tophi resolution, both of which were infrequent in this patient population; this indicated that studies with a larger sample size, longer follow-up or that recruit patients with higher baseline tophi are needed. Third, lesinurad's 400 mg daily dose in combination with allopurinol has a different safety profile than the 200 mg daily dose and placebo, characterised by higher rates of serious adverse events, sustained serum creatinine elevations and renal failure. Fourth, allopurinol alone even in average doses of 300 mg daily

leads to complete resolution of ≥1 target tophi in 29% of patients. This last message is sentinel, indicating that if we use appropriate doses of allopurinol, we are likely to not only improve gout management and resolve tophi, but also find a better use of newer therapies such as lesinurad and febuxostat and several more to come.

We need more data and longer follow-up studies of lesinurad (some of which are underway) and, in addition, studies that use optimal titrated doses of allopurinol at baseline that achieve target sUA. More comparative effectiveness studies comparing lesinurad to existing uricosurics will also help us better understand when best to use lesinurad versus probenecid or benzbromarone (where available). I hope that future lesinurad studies will use the therapeutic doses of allopurinol (800 mg/day in the USA; 900 mg/day in Europe), which are >300 mg daily in >50% of patients with gout, and not the suboptimal 'current standard' allopurinol 300 mg/day doses.

Competing interests JAS has received research grants from Takeda and Savient and consultant fees from Savient, Takeda, Regeneron, Merz, Iroko, Bioiberica, Crealta and Allergan pharmaceuticals, WebMD, UBM LLC and the American College of Rheumatology. JAS serves as the principal investigator for an investigator-initiated study funded by Horizon pharmaceuticals through a grant to DINORA, a 501 (c) (3) entity. JAS is a member of the executive of OMERACT, an organisation that develops outcome measures in rheumatology and receives arms-length funding from 36 companies; a member of the ACR's Annual Meeting Planning Committee; Chair of the ACR Meet-the-Professor, Workshop and Study Group Subcommittee; and a member of the Veterans Affairs Rheumatology Field Advisory Committee. JAS is supported by the resources and the use of facilities at the VA Medical Center at Birmingham, Alabama, USA.

Provenance and peer review Commissioned; externally peer reviewed.

To cite Singh JA. *Ann Rheum Dis* 2017;**76**:779–781.

Received 3 October 2016
Revised 12 November 2016
Accepted 8 December 2016
Published Online First 30 December 2016

► http://dx.doi.org/10.1136/annrheumdis-2016-209213

Ann Rheum Dis 2017;**76**:779–781.
doi:10.1136/annrheumdis-2016-210519

REFERENCES

1 Lopez AD, Williams TN, Levin A, *et al*. Remembering the forgotten non-communicable diseases. *BMC Med* 2014;12:200.
2 Zhu Y, Pandya BJ, Choi HK. Prevalence of gout and hyperuricemia in the US general population: The National Health and Nutrition Examination Survey 2007–2008. *Arthritis Rheumatol* 2011;63:3136–41.
3 Kuo CF, Grainge MJ, Zhang W, *et al*. Global epidemiology of gout: prevalence, incidence and risk factors. *Nat Rev Rheumatol* 2015;11:649–62.
4 Dehlin M, Drivelegka P, Sigurdardottir V, *et al*. Incidence and prevalence of gout in Western Sweden. *Arthritis Res Ther* 2016;18:164.
5 FDA News Release. FDA approves Zurampic to treat high blood uric acid levels associated with gout. 2015 (cited 10/01/2016). http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm478791.htm
6 European Medicines Agency. Zurampic. Lesinurad. 2016 (cited 10/01/2016). http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/003932/human_med_001963.jsp&mid=WC0b01ac058001d124
7 So A, Thorens B. Uric acid transport and disease. *J Clin Invest* 2010;120:1791–9.
8 Bardin T, Keenan RT, Khanna PP, *et al*. Lesinurad in combination with allopurinol: a randomised, double-blind, placebo-controlled study in gout patients with inadequate response to standard of care (the multinational CLEAR 2 study). *Ann Rheum Dis* 2017;76:811–20.
9 Saag KG, Fitz-Patrick D, Kopicko J, *et al*. Lesinurad combined with allopurinol: randomized, double-blind, placebo-controlled study in gout subjects with inadequate response to standard of care allopurinol (A US-based Study). *Arthritis Rheumatol* 2016. doi: 10.1002/art.39840. [Epub ahead of print 26 Aug 2016].
10 Sarawate CA, Patel PA, Schumacher HR, *et al*. Serum urate levels and gout flares: analysis from managed care data. *J Clin Rheumatol* 2006;12:61–5.
11 Singh JA, Hodges JS, Asch SM. Opportunities for improving medication use and monitoring in gout. *Ann Rheum Dis* 2009;68:1265–70.
12 Rees F, Jenkins W, Doherty M. Patients with gout adhere to curative treatment if informed appropriately: proof-of-concept observational study. *Ann Rheum Dis* 2013;72:826–30.
13 Becker MA, Schumacher HR Jr, Wortmann RL, *et al*. Febuxostat compared with allopurinol in patients with hyperuricemia and gout. *N Engl J Med* 2005;353:2450–61.
14 Higgins JP, Altman DG. Chapter 8: assessing risk of bias in included studies. In: Higgins JPT, Green S, ed. *Cochrane handbook for systematic reviews of interventions version 510 (updated March 2011)*. The Cochrane Collaboration, 2011. http://www.cochrane-handbook.org

2016 American College of Rheumatology/European League Against Rheumatism Criteria for Minimal, Moderate, and Major Clinical Response in Juvenile Dermatomyositis

An International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation Collaborative Initiative

Lisa G Rider,[1] Rohit Aggarwal,[2] Angela Pistorio,[3] Nastaran Bayat,[1] Brian Erman,[4] Brian M Feldman,[5] Adam M Huber,[6] Rolando Cimaz,[7] Rubén J Cuttica,[8] Sheila Knupp de Oliveira,[9] Carol B Lindsley,[10] Clarissa A Pilkington,[11] Marilynn Punaro,[12] Angelo Ravelli,[13] Ann M Reed,[14] Kelly Rouster-Stevens,[15] Annet van Royen-Kerkhof,[16] Frank Dressler,[17] Claudia Saad Magalhaes,[18] Tamás Constantin,[19] Joyce E Davidson,[20,21] Bo Magnusson,[22] Ricardo Russo,[23] Luca Villa,[24] Mariangela Rinaldi,[24] Howard Rockette,[2] Peter A Lachenbruch,[1] Frederick W Miller,[1] Jiri Vencovsky,[25] Nicolino Ruperto,[24] for the International Myositis Assessment and Clinical Studies Group and the Paediatric Rheumatology International Trials Organisation

Handling editor Tore K Kvien

For numbered affiliations see end of article.

Correspondence to
Dr Nicolino Ruperto, Istituto Giannina Gaslini, Pediatria II, PRINTO, Via G. Gaslini 5, Genoa 16147, Italy; nicolaruperto@gaslini.org.

LGR and RA contributed equally and JV and NR contributed equally.

This article is published simultaneously in the May 2017 issue of *Arthritis & Rheumatology*.

Submitted for publication 11 February 2016; accepted in revised form 31 January 2017

Accepted 1 March 2017
Published Online First 31 March 2017

ABSTRACT

To develop response criteria for juvenile dermatomyositis (DM). We analysed the performance of 312 definitions that used core set measures from either the International Myositis Assessment and Clinical Studies Group (IMACS) or the Paediatric Rheumatology International Trials Organisation (PRINTO) and were derived from natural history data and a conjoint analysis survey. They were further validated using data from the PRINTO trial of prednisone alone compared to prednisone with methotrexate or cyclosporine and the Rituximab in Myositis (RIM) trial. At a consensus conference, experts considered 14 top candidate criteria based on their performance characteristics and clinical face validity, using nominal group technique. Consensus was reached for a conjoint analysis–based continuous model with a total improvement score of 0–100, using absolute per cent change in core set measures of minimal (≥30), moderate (≥45), and major (≥70) improvement. The same criteria were chosen for adult DM/polymyositis, with differing thresholds for improvement. The sensitivity and specificity were 89% and 91–98% for minimal improvement, 92–94% and 94–99% for moderate improvement, and 91–98% and 85–86% for major improvement, respectively, in juvenile DM patient cohorts using the IMACS and PRINTO core set measures. These criteria were validated in the PRINTO trial for differentiating between treatment arms for minimal and moderate improvement (p=0.009–0.057) and in the RIM trial for significantly differentiating the physician's rating for improvement (p<0.006). The response criteria for juvenile DM consisted of a conjoint analysis–based model using a continuous improvement score based on absolute per cent change in core set measures, with thresholds for minimal, moderate, and major improvement.

Juvenile dermatomyositis (DM) is a systemic autoimmune disease characterised by chronic skeletal muscle inflammation and weakness. Core set measures to assess juvenile DM disease activity have been established and validated by the International Myositis Assessment and Clinical Studies Group (IMACS) and the Paediatric Rheumatology International Trials Organisation (PRINTO), with provisional endorsement by the American College of Rheumatology and the European League Against

To cite: Rider LG, Aggarwal R, Pistorio A, *et al*. *Ann Rheum Dis* 2017;**76**:782–791.

This criteria set has been approved by the American College of Rheumatology (ACR) Board of Directors and the European League Against Rheumatism (EULAR) Executive Committee. This signifies that the criteria set has been quantitatively validated using patient data, and it has undergone validation based on an independent data set. All ACR/EULAR-approved criteria sets are expected to undergo intermittent updates.

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

Rheumatism.[1–6] Both core sets include physician and parent global activity, muscle strength, and physical function. IMACS also includes the most abnormal serum muscle enzyme value and extramuscular global activity, whereas PRINTO includes instead a health-related quality of life measure, the Child Health Questionnaire[7] and a global activity score, the Disease Activity Score.[8] IMACS measures muscle strength using manual muscle testing, and PRINTO measures muscle strength using the Childhood Myositis Assessment Scale.[1 2 5]

Combinations of these measures to determine clinical improvement were developed to enhance the sensitivity of responses and decrease the sample sizes needed, by using large prospective natural history data sets and expert clinician consensus as the gold standard. For both PRINTO and IMACS, at least 20% improvement in 3 of 6 core set measures with no more than 1 or 2 worsening (which cannot be muscle strength) had been established as preliminary response criteria, and additional combinations of improvement in the core set measures serve as secondary response criteria.[9 10] PRINTO adapted its top criteria for minimal clinical improvement to moderate and major improvement by using cutoffs of 50% and 70%, similar to the improvement criteria for juvenile idiopathic arthritis (JIA).[11–13]

Although the preliminary response criteria for juvenile DM advanced the assessment of patients and their responses to treatment, those criteria were limited by differences in the core set measures and final consensus response criteria between IMACS and PRINTO, a lack of randomised controlled trial data for full validation, and inadequate exploration of more sensitive approaches using hybrid or continuous methods.[14] The preliminary response criteria also considered each core set measure equally rather than differentially weighting them. However, most myositis experts agree that some core set measures are more important, such as physician global activity and muscle strength.[3 15] For PRINTO studies, physician global evaluation of disease activity, muscle strength, and parent global evaluation of the child's overall well-being were weighted as the most important core set measures in a logistic regression analysis.[3 10] Moreover, the preliminary response criteria did not validate criteria for moderate or major improvement. There is, therefore, a clear need to have standardised improvement criteria for all levels of improvement in future clinical trials, similar to the standardized criteria developed for rheumatoid arthritis (RA) and JIA.

For these reasons, IMACS and PRINTO engaged in a joint effort to develop fully validated response criteria for juvenile DM, including criteria for minimal, moderate, and major clinical response. This report focuses on the consensus conference in which the top candidate definitions of response leading to the final juvenile DM response criteria were considered.

METHODS

In previous reports,[16 17] we described the methodology used a) to create patient profiles using natural history data and obtain expert consensus on minimal, moderate, and major improvement,[16] b) to determine differential weights of the core set measures using conjoint analysis, and c) to draft six types of candidate definitions for response criteria using the myositis expert survey on thresholds of improvement and data-driven methods, such as logistic regression and conjoint analysis (table 1).

Conjoint analysis is a choice modeling or discrete choice experiment, which is a valid methodology for developing composite criteria and has been used recently in rheumatology.[19–22] In the conjoint analysis surveys administered using 1000Minds online software,[23] experts were presented with pairs of hypothetical patient scenarios; each patient had different levels of improvement in the same 2 core set measures, assuming other core set measures remained the same. Experts rated which of the 2 scenarios had greater improvement. Based on the rater's response, the relative weights of core set measures and their levels of improvement were established and used to develop a scoring system by mathematical methods based on linear programming[24] such that when all 6 core set measures are considered together, the maximum score (total improvement score) possible for representing a patient's improvement is 100, and the minimum score is 0.

We then compared the performance characteristics of the drafted definitions in the patient profiles, using expert consensus ratings as a gold standard, and externally validated the candidate response criteria by applying them to clinical trial data. This process led to the development of traditional categorical as well as continuous candidate definitions for response criteria, with thresholds for minimal, moderate, and major improvement.[18] Continuous candidate definitions can also be considered hybrid definitions, because the same definition can be used either as a continuous outcome measure by using the total improvement score or as a categorical outcome measure by using the thresholds for minimal, moderate, and major improvement.

Candidate definitions were evaluated using consensus profile ratings as the gold standard, by assessing sensitivity, specificity, and area under the curve (AUC) to compare the performance of these candidate definitions. Those that performed well in the consensus profiles (sensitivity and specificity both ≥80%, AUC ≥0.9 for minimal, and AUC ≥0.8 for moderate and major improvement, using IMACS or PRINTO core set measures[1]) were externally validated. The PRINTO trial randomised patients with new-onset juvenile DM to receive prednisone alone (n=47) or prednisone combined with methotrexate or cyclosporine (n=46 patients per treatment arm).[13] χ^2 analysis was used to compare the percentages of patients meeting the candidate definitions for response at the primary end point (6 months) for the combined treatment arms versus the prednisone-alone (placebo) arm. Definitions with a significant difference ($p<0.05$) between treatment arms for minimal improvement were further considered. Both PRINTO and IMACS core set measures were available in this trial.

A second trial validation data set included 48 juvenile DM patients enrolled in the Rituximab in Myositis (RIM) trial for treatment-refractory patients. It had a randomised placebo-phase design in which patients received either rituximab or placebo at weeks 0 and 1, and at weeks 8 and 9 their treatment assignment was reversed in a blinded manner.[25] We used the Mann-Whitney U test to determine whether each candidate definition could differentiate between the treating physician's rating of improvement (score range 1–7) at 6 months, a time point when most patients improved and that was also comparable to that in the PRINTO trial. For the RIM trial, only the IMACS core set measures were available.

We then selected the top candidate definitions, up to 4 top-performing definitions from each of the six different types of candidate definitions (table 1), for consideration at the final consensus conference as a manageable number of definitions to discuss.

Consensus conference

Nominal group technique was used at a consensus conference held in Paris, France on 9–10 June 2014, led by experienced moderators (LGR and NR, for the paediatric working group) The methodologies used to develop the new candidate response criteria and performance characteristics of each type of candidate definition were reviewed with the participants in a general

Table 1 Types of candidate definitions for response criteria that were developed and tested

Type of candidate definitions of response	Description	Example of the candidate definition for the response criteria
Previously published (categorical definition)	Previously published response criteria that were retested	Minimal. Three of any 6 improved by ≥20%, no more than 1 worse by >30% (which cannot be CMAS)[10] Moderate. Three of any 6 improved by ≥50%, no more than 1 worse by >30% (which cannot be CMAS)[11] Major. Three of any 6 improved by ≥70%, no more than 1 worse by >30% (which cannot be CMAS)[11]
Newly drafted (categorical definition)	Drafted relative or absolute per cent change in candidate definitions of response, based on recent CSM survey	Minimal. MD global, muscle strength (MMT or CMAS), and 1 other CSM improved by ≥20% Moderate. MD global, muscle strength (MMT or CMAS), and 1 other CSM improved by ≥30% Major. MD global, muscle strength (MMT or CMAS), and 1 other CSM improved by ≥50%
Weighted (categorical definition)	Applied conjoint analysis relative weights to CSM in newly drafted definitions; each CSM receives improvement points (corresponding relative weights) when it reaches the threshold for minimal, moderate, or major improvement; worsening points are applied similarly; improvement is calculated based on a total score of improvement vs worsening	Improvement=at least 3.5 improvement points of 10 total improvement points, and no more than 1.5 worsening points, where MD global=2 points, parent global=1 point, MMT/CMAS=3 points, C-HAQ=1.5 points, extramusc/DAS=1.5 points, enzyme/CHQ-PhS=1 point Minimal. Improvement points given when CSM ≥20%; worsening points given when CSM worse by >30% Moderate. Improvement points given when CSM ≥50%; worsening points given when CSM worse by >30% Major. Improvement points given when CSM ≥75%; worsening points given when CSM worse by >30%
Logistic regression (continuous definition)	Model of improvement using a combination of CSM with different weights, as developed in the logistic regression model; total scores derived, with different cutoffs for minimal, moderate, and major improvement Relative % change	Improvement score=(MD global % change)+0.5×(parent global activity % change)+0.5×(extramusc activity or DAS % change) Minimal. Improvement score ≥15 Moderate. Improvement score ≥30 Major. Improvement score ≥60
CSM–weighted (continuous definition)*	Multiply the % change in each CSM by the weights derived from conjoint analysis, then sum (% change in each CSM×conjoint analysis weights) to get final total improvement score; different thresholds for minimal, moderate, and major improvement established based on consensus profile ratings as gold standard	Improvement score=2×(MD global % change)+(parent global % change)+3×(MMT or CMAS % change)+1.5×(C-HAQ % change) +1.5×(extramusc or DAS % change)+(enzyme or CHQ-PhS % change) Minimal. Improvement score ≥100 Moderate. Improvement score ≥250 Major. Improvement score ≥400
Conjoint analysis (continuous definition)	For a given range in the level of improvement in each CSM, a score is assigned, as developed by the survey results and modelling; greater degrees of improvement receive higher scores; a patient is minimally improved if the improvement score is above the cutoff for minimal improvement; similarly for moderate and major improvement	Cut points for the model for juvenile DM are: Minimal. Improvement score ≥30 Moderate. Improvement score ≥45 Major. Improvement score ≥70†

*This type of definition was not brought to the final consensus conference.
†The full absolute per cent change model is shown in table 3 and in online supplementary table S2 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40060/abstract).
C-HAQ, Childhood Health Assessment Questionnaire; CHQ-PhS, physical summary score of the Child Health Questionnaire–Parent Form 50; CMAS, Childhood Myositis Assessment Scale; CSM, core set measure; DAS, Disease Activity Score; DM, dermatomyositis; enzyme, most abnormal serum muscle enzyme value among aldolase, alanine aminotransferase, aspartate aminotransferase, lactate dehydrogenase, and creatine kinase; extramusc, extramuscular global activity; MD global, physician global activity score; MMT, manual muscle testing; parent global, parent global activity score.

session. The 12 paediatric working group participants first independently and then as a group reviewed the performance characteristics of the 14 top candidate definitions of response criteria for juvenile DM. Data for minimal, moderate, and major clinical response were presented for each definition, including a detailed spreadsheet that included the performance in the patient profiles using the IMACS and PRINTO core set measures, including sensitivity, specificity, AUC, as well as kappa values and ORs. AUC was defined as the average of the sensitivity and specificity values for all categorical candidate definitions, as well as for thresholds for minimal, moderate, and major improvement in continuous candidate definitions. In addition, for continuous definitions, an AUC for the total improvement score was determined from the receiver operating characteristic (ROC) curve as a plot of sensitivity versus (1—specificity) for total improvement scores as well as for thresholds.[26–28] Results of the external validation for each candidate definition from the PRINTO and RIM clinical trial data sets were also presented.

Paediatric working group

After reviewing the performance of the 14 top performing candidate definitions, the 12 paediatric working group participants developed consensus response criteria for minimal, moderate, and major improvement in juvenile DM. The participants were informed of the secondary goal of reaching consensus on response criteria for both juvenile DM and adult DM/polymyositis (PM). Participants were first asked to rank their top five choices, considering the data presented, based on face validity, feasibility, and generalisability, and to determine which response criteria were most clinically meaningful. The voting process was conducted in a systematic manner with a predetermined

format using nominal group technique[29 30] facilitated by an internet-based system developed by staff at the PRINTO coordinating centre.[31 32] Voting was done anonymously and independently using the online voting software.

After the initial round of voting, the results were shared with the group. Each participant was then asked to explain his or her top- and bottom-ranked choices to the group. The rounds of voting continued in the same manner until consensus was reached (≥80% of the votes) or until it was clear that consensus would not be reached. Between each round, after the participants were shown the results, the administrators were allowed to remove candidate definitions that decisively received a small proportion of the votes. In the final round, participants were asked to select their final top response criteria. The paediatric working group also voted on additional issues, including use of both IMACS and PRINTO core set measures and response criteria for juvenile DM that would interchange both the IMACS and PRINTO measures. Participants also voted on retesting the performance of the top candidate response criteria in future trials.

Combined paediatric and adult working group

After consensus was attained for juvenile DM response criteria, a combined working group of 22 paediatric and adult experts was formed to determine whether consensus could be reached on final, common response criteria for both juvenile DM and adult DM/PM. Common response criteria that would include both juvenile DM and adult DM/PM were considered for use in clinical trials, which might facilitate drug approvals for myositis treatment. Experienced moderators (LGR, RA, FWM, and NR) led the combined working group. For the first round of votes, the top adult and paediatric definitions from the final round of voting in each working group were considered. The online voting system was utilised again, and each participant discussed his or her top-choice candidate definition, using nominal group technique in a round-robin manner. At each round, participants were asked to select only one candidate top response croteroa set; discussion was stopped once consensus of ≥80% was reached. For determining the thresholds of improvement for the selected definition, the required consensus was ≥70%, which was done by post-conference voting.

RESULTS

The performance characteristics of 101 of 312 candidate definitions were excellent (sensitivity and specificity of ≥80%, AUC ≥0.90 for minimal improvement), and 30 candidate definitions also performed well in 2 clinical trials, in which they differentiated between treatment arms ($p<0.05$ for minimal improvement) and differentiated the treating physician's improvement score at week 24 ($p<0.001$).[15]

Top candidate definitions for response criteria

Fourteen top-performing candidate definitions were brought to the paediatric working group for consideration at the consensus conference (table 2 and online supplementary tables S1 and S2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40060/abstract). These candidate criteria included nine categorical definitions in which different criteria were set for minimal, moderate, and major improvement and five continuous definitions in which improvement points are given on a continuous scale that corresponds to the magnitude of improvement, with different thresholds for minimal, moderate, and major improvement. Among the nine categorical definitions, two were previously published IMACS and PRINTO response criteria,[9–11] four were newly drafted definitions based on a survey of experts, and three were weighted definitions. Among the continuous definitions, two were developed by logistic regression, and three were developed from the conjoint analysis survey. Among the 14 candidate criteria considered, 11 were based on relative per cent change, and 3 were based on absolute per cent change in the core set measures.

The performance characteristics of these 14 candidate definitions are shown in table 2 and online supplementary table S1 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40060/abstract). In the patient profiles, with expert consensus as a gold standard, all definitions presented at the conference had sensitivity and specificity of ≥87% (AUC≥0.90) for minimal improvement (table 2 and online supplementary table S1). For moderate improvement, specificity decreased but was ≥80% (AUC≥0.88), and for major improvement specificity was generally ≥75% (AUC≥0.84). For continuous definitions, the AUCs (from ROC curves) for the total improvement score were generally better than the AUCs (average of sensitivity and specificity) for the thresholds of minimal, moderate, and major improvement. Performance was similar between the IMACS and PRINTO core set measures for each definition.

Almost all candidate criteria were validated using the PRINTO trial at 6 months, when they could differentiate between treatment arms, with $p<0.05$ for minimal improvement (table 2 and online supplementary table S1). All candidate criteria were also validated in 48 juvenile DM patients in the RIM trial.[25] All definitions could differentiate the median treating physician's improvement score at week 24 ($p\leq0.006$).

Consensus conference voting

Among the 14 candidate definitions, 13 and 11 candidate definitions of response were promoted in the first and second voting rounds, respectively. In round three, six candidate definitions were chosen, each receiving a similar number of votes. These six included the three conjoint analysis–based continuous definitions, a conjoint analysis–based weighted definition, a logistic regression absolute per cent change definition, and the previously published PRINTO preliminary response criteria.[8 9] In the fourth round of voting and discussion, participants reached consensus on final top response criteria, a conjoint analysis–based continuous model using absolute per cent change in the IMACS or PRINTO core set measures (table 3).

Table 2 and online supplementary table S1 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40060/abstract) show the performance characteristics in the patient profiles and the trial validation for each of the top candidate response criteria presented at the conference. For the top conjoint analysis–based continuous response criteria using absolute per cent change in each of the core set measures, the sensitivity and specificity in the patient profiles was generally >90% and the AUC >0.90 for both the IMACS and PRINTO measures. For the PRINTO trial, a difference in the treatment arms was detected for minimal and moderate improvement using the top response criteria, and in the RIM trial a difference in the physician's rating of improvement when the response criteria rated the patient as improved versus not improved was detected for minimal, moderate, and major improvement.

Paediatric experts favoured the conjoint analysis–based continuous response criteria because of the continuous improvement score that corresponds to the magnitude of improvement and provides the ability to categorise a patient's degree of change into minimal, moderate, and major improvement. The continuous model definitions also differentially weight the

Table 2 Detailed performance characteristics of patient profiles for the top 5 candidate definitions presented at the consensus conference*

Candidate definition type based on final consensus rank order, improvement category, CSM	Sensitivity, %	Specificity, %	Threshold AUC†	Total improvement score AUC‡	PRINTO trial§ Tx (%)	PRINTO trial§ Ctrl (%)	PRINTO trial§ *p* Value	RIM trial¶ Response criteria, improved**	RIM trial¶ Response criteria, not improved#	RIM trial¶ *p* Value	Rank
Conjoint analysis, absolute % change (model 3)††											1
Minimal (≥30)											
IMACS	89	91	0.90	0.98	75	53	0.009	2.0	3.0	<0.001	
PRINTO	89	98	0.93	0.99	73	55	0.038				
Moderate (≥45)											
IMACS	92	99	0.95	0.99	70	53	0.057	2.0	3.0	<0.001	
PRINTO	94	94	0.94	0.99	71	51	0.023				
Major (≥70)											
IMACS	91	86	0.89	0.96	51	43	0.341	2.0	3.0	0.006	
PRINTO	98	85	0.91	0.98	58	49	0.331				
Conjoint analysis, relative % change (model 1)‡‡											2
Minimal (≥33)											
IMACS	99	87	0.93	0.98	75	55	0.018	2.0	4.0		
PRINTO	96	98	0.97	1.00	74	55	0.027				
Moderate (≥60)											
IMACS	97	93	0.95	0.99	73	51	0.011	2.0	3.0	<0.001	
PRINTO	97	96	0.96	1.00	70	51	0.032				
Major (≥80)											
IMACS	91	87	0.89	0.96	57	49	0.396	1.5	3.0	<0.001	
PRINTO	98	86	0.92	0.97	61	49	0.179				
Conjoint analysis, relative % change (model 2)‡‡											3
Miminal (≥33)											
IMACS	95	94	0.94	0.98	75	53	0.009	2.0	4.0	<0.001	
PRINTO	94	98	0.96	0.99	74	55	0.027				
Moderate (≥55)											
IMACS	95	95	0.95	1.00	70	51	0.032	2.0	3.0	<0.001	
PRINTO	97	98	0.98	1.00	70	51	0.032				
Major (≥77)											
IMACS	93	86	0.90	0.97	49	47	0.814	1.0	2.0	0.011	
PRINTO	96	90	0.93	0.99	59	49	0.273				
Weighted definition, relative % change§§											4
Minimal (improvement points given when CSM ≥20, worsening points given when CSM worse by >30)											
IMACS	95	100	0.97	NA	70	51	0.032	2.0	3.0	<0.001	
PRINTO	92	98	0.95	NA	73	53	0.021				
Moderate (improvement points given when CSM≥50%, worsening points given when CSM worse by >30%)											
IMACS	95	91	0.93	NA	68	51	0.045	2.0	3.0	<0.001	
PRINTO	95	92	0.94	NA	71	51	0.023				
Major (improvement points given when CSM ≥75%, worsening points given when CSM worse by >30%)											
IMACS	100	81	0.91	NA	64	47	0.050	1.5	3.0	<0.001	
PRINTO	98	85	0.91	NA	62	49	0.142				

Continued

Table 2 Continued

Candidate definition type based on final consensus rank order, improvement category, CSM	Sensitivity, %	Specificity, %	Threshold AUC†	Total improvement score AUC‡	PRINTO trial§ Tx (%)	PRINTO trial§ Ctrl (%)	PRINTO trial§ *p* Value	RIM trial¶ Response criteria, improved**	RIM trial¶ Response criteria, not improved#	RIM trial¶ *p* Value	Rank
Previously published definition,[10 11] relative % change											5
Minimal (3 of any 6 improved by ≥20%, no more than 1 worse by >30%) (which cannot be MMT/CMAS)[10]											
IMACS	93	100	0.97	NA	70	51	0.032	2.0	3.0	<0.001	
PRINTO	88	100	0.94	NA	71	51	0.023				
Moderate (3 of any 6 improved by ≥50%, no more than 1 worse by >30%) (which cannot be MMT/CMAS)[11]											
IMACS	90	95	0.93	NA	66	51	0.081	2.0	3.0	<0.001	
PRINTO	90	96	0.93	NA	68	51	0.045				
Major (3 of any 6 improved by ≥70%, no more than 1 worse by >30%) (which cannot be MMT/CMAS)[11]											
IMACS	99	83	0.91	NA	63	49	0.111	2.0	3.0	<0.001	
PRINTO	99	89	0.94	NA	60	49	0.223				

*The performance characteristics of patient profiles for definitions ranked 6–14 are shown in online supplementary table S1 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40060/abstract). Note that either International Myositis Assessment and Clinical Studies (IMACS) or Paediatric Rheumatology International Trials Organisation (PRINTO) core set measures (CSMs) may be used in these candidate definitions of response; the candidate definitions were developed in parallel with IMACS or PRINTO CSMs. Tx, treatment arm of prednisone in combination with methotrexate or cyclosporine; Ctrl, control; NA, not applicable.
†Calculated as the area under the curve (AUC) from the receiver operating characteristic (ROC) curve for the total improvement score and the threshold for minimal, moderate, and major improvement.
‡Calculated as the AUC from the ROC curve, using the total improvement score and the threshold cutoffs for minimal, moderate, and major improvement, which applies only to continuous definitions.
§PRINTO juvenile dermatomyositis (DM) trial of prednisone alone versus prednisone with methotrexate or cyclosporine (n=139).[13]
¶Rituximab in Myositis (RIM) trial juvenile DM arm (n=48). Comparison of the treating physician's rating of improvement if the improvement criteria are met versus not met at week 24.[25] A 1-point difference in physician's rating of improvement from no improvement to minimal improvement was considered not only statistically significant but also clinically significant.
**Median score for physician's rating of improvement.
††The conjoint analysis–based continuous candidate response criteria using absolute per cent change in CSMs (absolute per cent change model) are shown in table 3. These criteria are also the top response criteria for adult DM/polymyositis (PM), but with different thresholds for the total improvement score for minimal, moderate, and major improvement.[18]
‡‡The conjoint analysis–based continuous candidate definitions using relative per cent change in CSMs are shown in online supplementary table S3 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40060/abstract). These criteria are also the second- and third-choice criteria for adult DM/PM, but with different thresholds in the total improvement score for minimal, moderate, and major improvement.[18]
§§Improvement=at least 3.5 improvement points of 10 total improvement points, and no more than 1.5 worsening points, where physician global activity=2 points, parent global activity=1 point, manual muscle testing (MMT) or Childhood Myositis Assessment Scale (CMAS)=3 points, Childhood Health Assessment Questionnaire=1.5 points, extramuscular global activity or Disease Activity Score=1.5 points, and enzyme or physical summary score of the Child Health Questionnaire–Parent Form 50=1 point.

various core set measures, which experts thought were consistent with their assessment of the relative importance of each of the core set measures. The top response criteria were based on absolute per cent change in core set measures, which was also favoured by the participants because, given the various visual analog scale (VAS) measurements used in the core set measures, the absolute per cent changes were more congruent than relative per cent changes with actual changes that the myositis experts see in clinical practice.

Final response criteria chosen by the combined pediatric and adult working group

For this round of votes, the top 2 paediatric (table 2) and adult definitions[18] were considered. Two rounds of voting resulted in final consensus response criteria, with 91% of participants voting for the conjoint analysis–based continuous response criteria based on absolute per cent change in the core set measures (table 3). It was agreed that the top response criteria would be used in future clinical trials that combined juvenile DM and adult DM/PM. Because the final response criteria were similar, participants favoured using response criteria that would be common to juvenile DM and adult DM/PM, and they favoured combined studies when possible as well as the possibility of comparing outcomes in separate studies using the same final response criteria.

Other votes

In a post-conference final vote using the Delphi method, 74% of the participants agreed to use the following paediatric threshold values for minimal, moderate, and major response in juvenile DM: total improvement score ≥30 (on a scale of 0–100) for minimal, ≥45 for moderate, and ≥70 for major improvement. In contrast, the final thresholds for minimal, moderate, and major response in adult DM/PM were ≥20, ≥40, and ≥60, respectively. The paediatric working group also reached consensus that, given the overall similarity between the IMACS and PRINTO response criteria, joint IMACS/PRINTO response criteria for juvenile DM are being proposed. The current development of the response criteria in parallel between the IMACS and PRINTO core set measures necessitates that either all of the IMACS or all of the PRINTO core set measures be used. The paediatric experts, however, committed to measure both IMACS and PRINTO core set measures in future therapeutic trials, with

Table 3 Final top response criteria for minimal, moderate, and major improvement in juvenile dermatomyositis (DM) and combined adult DM/PM and juvenile DM clinical trials and studies*

Core set measure, level of improvement based on absolute per cent change	Improvement score
Physician global activity	
Worsening to 5% improvement	0
>5% to 15% improvement	7.5
>15% to 25% improvement	15
>25% to 40% improvement	17.5
>40% improvement	20
Parent global activity	
Worsening to 5% improvement	0
>5% to 15% improvement	2.5
>15% to 25% improvement	5
>25% to 40% improvement	7.5
>40% improvement	10
Manual muscle testing or CMAS	
Worsening to 2% improvement	0
>2% to 10% improvement	10
>10% to 20% improvement	20
>20% to 30% improvement	27.5
>30% improvement	32.5
Childhood Health Assessment Questionnaire	
Worsening to 5% improvement	0
>5% to 15% improvement	5
>15% to 25% improvement	7.5
>25% to 40% improvement	7.5
>40% improvement	10
Enzyme (most abnormal) or CHQ-PhS	
Worsening to 5% improvement	0
>5% to 15% improvement	2.5
>15% to 25% improvement	5
>25% to 40% improvement	7.5
>40% improvement	7.5
Extramuscular activity or Disease Activity Score	
Worsening to 5% improvement	0
>5% to 15% improvement	7.5
>15% to 25% improvement	12.5
>25% to 40% improvement	15
>40% improvement	20

The **total improvement score** is the sum of all 6 improvement scores associated with the change in each core set measure. A total improvement score of **≥30** represents **minimal improvement**, a score of **≥45** represents **moderate improvement**, and a score of **≥70** represents **major improvement**.

*Either all of the International Myositis Assessment and Clinical Studies Group (IMACS) or all of the Paediatric Rheumatology International Trials Organisation (PRINTO) core set measures may be used. Note that these response criteria are also proposed for use in combined adult dermatomyositis/polymyositis (DM/PM) and juvenile DM trials.[18] For comparison, the thresholds of improvement in the total improvement score for adult DM/PM are ≥20 for minimal improvement, ≥40 for moderate improvement, and ≥60 for major improvement.
How to calculate the improvement score: The absolute percent change ([final value − baseline value]/range × 100) is calculated for each core set measure. For muscle enzymes, the most abnormal serum muscle enzyme level at baseline (creatine kinase, aldolase, alanine transaminase, aspartate aminotransferase, lactate dehydrogenase) is used. The enzyme range was calculated based on a 90% range of enzymes from natural history data,[5 38] which for creatine kinase is 15 times the upper limit of normal (ULN), for aldolase is 6 times the ULN, and for lactate dehydrogenase, aspartate aminotransferase, and alanine transaminase is 3 times the ULN. The ULN is determined according to the individual laboratories in the participating centers. The ranges for the other core set activity measures are based on the instrument scale used.[13 15 25] An improvement score is assigned for each core set measure based on the absolute percent change. These are totaled among the 6 IMACS or PRINTO core set measures. The thresholds for minimal, moderate, and major improvement are provided. The total improvement score itself may also be compared among treatment arms in a trial. A total improvement score between 0 and 100 corresponds to the degree of improvement, with higher scores corresponding to a greater degree of improvement.
CHQ-PhS, Physical Summary Score of the Child Health Questionnaire–Parent Form 50; CMAS, Childhood Myositis Assessment Scale; DAS, Disease Activity Score; MMT, manual muscle testing.

92% agreement, and to continue to test the interchangeability of the IMACS and PRINTO core set measures. The group also unanimously agreed to retest the validity of the top five candidate definitions for response criteria and to utilise the other four definitions as secondary end points in future clinical trials. The top 3 of these criteria, the conjoint analysis definitions, are the same for both juvenile DM and adult DM/PM, with different thresholds of improvement (table 3 and online supplementary table S3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40060/abstract).

DISCUSSION

Conjoint analysis–based continuous response criteria, based on absolute per cent change in the core set measures, were

developed as the consensus- and data-driven response criteria for minimal, moderate, and major improvement in juvenile DM. For the response criteria, either IMACS or PRINTO core set measures could be used. In addition, it was agreed that the same response criteria, using the IMACS core set measures but with different thresholds for improvement, would be the consensus response criteria for adult DM/PM trials and combined juvenile DM and adult DM/PM trials in the future.[18]

The comprehensive process used to develop final response criteria for minimal, moderate, and major improvement in juvenile DM included the use of large, prospective, natural history data sets for juvenile DM and data from two randomised controlled trials for validation, which included a wide range of disease activity and different stages of disease, from recently diagnosed to treatment-refractory patients.[13 15 25] The involvement of many clinical experts who had experience using the core set measures in juvenile DM patients was also critical. They provided input at several points throughout the process, including determining thresholds for improvement in core set measures by which definitions of response were drafted, achieving gold standard ratings of improvement by evaluating and developing consensus patient profiles, completing the conjoint analysis surveys to develop differential weights for the core set measures, and participating in the final consensus conference to achieve consensus for common response criteria with the greatest clinical face validity. The current response criteria (table 3) also resolve the differences between PRINTO and IMACS core set measures by testing candidate definitions of response criteria in parallel using both sets of measures and showing that they are largely interchangeable, and that their performance is comparable. Moreover, this project brought both IMACS and PRINTO consortia to work together for this rare disease.

The combined group of paediatric and adult experts selected the same top-choice definition but with differing thresholds for improvement, which had very similar performance characteristics and were thought to be more appropriate for use in clinical trials that would, in the future, combine adult and paediatric patients.

The final response criteria selected, conjoint analysis–based continuous response criteria using absolute per cent change in core set measures, have many advantages. For each measure, improvement points are calculated based on the level of change in that measure, and each core set measure is differentially weighted, such that changes in muscle strength and physician global activity are weighted more heavily than changes in the most abnormal enzyme value or quality of life. A total improvement score can be obtained as a continuous measure, and the means or medians of total improvement scores can be compared between treatment arms.[33] A total improvement score between 0 and 100 also corresponds to the degree of improvement, with higher scores corresponding to a greater magnitude of improvement. This score may be more sensitive to change, resulting in smaller trial sample sizes.[33 34] Alternatively, thresholds for minimal, moderate, and major improvement have been established that allow dichotomous use of the response criteria as well. Therefore, this is truly a hybrid model that can be used as either a continuous or categorical outcome measure within the same response criteria depending on the trial design and needs of the study.

The response criteria allow input from all the core set measures instead of relying on only a few measures to determine whether a patient has experienced improvement. However, although these response criteria were developed using all six core set measures, the response criteria could still be used if fewer core set measures were obtained, allowing for greater flexibility in the types of patients and improvements that can occur, but we caution that the response criteria are most accurate when all six core set measures are used. As such, the response criteria signify a major advance in assessing improvement in therapeutic trials and other clinical research studies by providing data-driven response criteria that were developed by consensus of major stakeholders in the field who come from all over the world.

Prior response criteria in rheumatic diseases have included[33 34] relative per cent change,[35 36] whereas myositis response criteria are based on absolute per cent change. The experts favoured the use of absolute per cent change for various reasons. In this study, several core set measures used a 10-cm VAS, and the experts thought that absolute per cent change better represents the degree of change they see in clinical practice. Moreover, absolute per cent changes can be calculated when the baseline core set measure is 0 and give similar results for similar degrees of change at either end of the VAS.

The participants also favoured using the same response criteria for juvenile DM and adult DM/PM, but with cut points or thresholds for improvement specific to paediatric or adult patients. Having common response criteria facilitates the potential to conduct combined clinical trials, such as the RIM trial,[25] and to compare the outcomes of trials and studies conducted separately. Participants agreed to include other top-performing definitions that were highly rated as secondary end points for future clinical trials. Among these were not only other conjoint analysis–based continuous models but also the published PRINTO preliminary response criteria.[10 11] Future work should also evaluate whether a baseline composite score threshold derived from the PRINTO or IMACS core set measures could be used as inclusion criteria for future clinical trials.

Limitations of the present work include the lack of a placebo group in the RIM trial. For this reason, the physician's assessment of improvement at 6 months was used instead. We were fortunate to have another controlled clinical trial for juvenile DM that had three treatment arms to use for external validation,[13] in which we evaluated the ability of the candidate definitions to differentiate between treatment arms. Although thresholds for major improvement were developed and validated in fewer patients, we believe that it was sufficient given that 29% of patients had major improvement in patient profiles, and 17% had major improvement in the clinical trials used for validation. The final conjoint analysis–based continuous response criteria also do not address worsening in the core set measures; however, this generally does not affect the outcome, because when patients are rated as improved, no more than 1 or 2 measures worsen in our clinical data sets. Also, although we tested the interchange of IMACS and PRINTO core set measures, we tested these variations as 2 parallel core set measures but did not examine intermixing the PRINTO and IMACS core set measures. Further work to examine the interchangeability of the IMACS and PRINTO core set measures will be needed.

The data sets used to develop the new response criteria primarily contained information about patients with a recent diagnosis or those experiencing a disease flare, and further work is needed to determine how the response criteria perform in patients with longstanding disease or those with significant disease-related damage. Finally, although application of the criteria might seem cumbersome, as regularly done for JIA and RA, the evaluation of improvement will be facilitated by appropriate dedicated software or 'apps', or in the future, by simplification of the manner in which the core set measures are evaluated (eg, similar to the

Juvenile Arthritis Disease Activity Score for JIA).[37] The time required to apply these criteria is estimated to be 25–35 min to complete the core set measures at each visit[1] and 2–3 min to hand-calculate the total improvement score and degree of response. Both IMACS and PRINTO are developing a web-based tool as well as a downloadable calculator that will allow easy administration of the response criteria and immediate calculation. The apparent complexity is, however, counterbalanced by the establishment of different validated levels of improvement, which constitute the real novelty of this project and which have never been validated as such for either RA or JIA, despite being regularly reported in clinical trials.

In summary, conjoint analysis–based continuous response criteria that establish different thresholds for minimal, moderate, and major improvement and utilise the absolute per cent change in core set measures were chosen as the consensus response criteria for juvenile DM and were validated using both natural history and trial data. These response criteria should be highly acceptable and widely used given that they were developed with consensus among many myositis experts worldwide. They should be sensitive in detecting differences in improvement and in quantitating the degree of improvement, as seen in the two clinical trials. Thus, clinical trials that test new therapies for juvenile DM should be easier to design, conduct, and compare.

Author affiliations
[1]NIEHS, NIH, Bethesda, Maryland, USA
[2]University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[3]Istituto Giannina Gaslini, Servizio di Epidemiologia e Biostatistica, Genoa, Italy
[4]Social and Scientific Systems, Inc., Durham, North Carolina, USA
[5]The Hospital for Sick Children, Toronto, Ontario, Canada
[6]IWK Health Centre, Halifax, Nova Scotia, Canada
[7]University of Firenze, Florence, Italy
[8]Hospital de Niños Pedro de Elizalde, University of Buenos Aires, Buenos Aires, Argentina
[9]Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[10]University of Kansas City Medical Center, Kansas City, Kansas, USA
[11]Great Ormond Street Hospital for Children NHS Trust, London, UK
[12]University of Texas Southwestern Medical Center, Dallas, Texas, USA
[13]Istituto Giannina Gaslini, Pediatria II - Reumatologia, and Università degli Studi di Genova, Genoa, Italy
[14]Duke University, Durham, North Carolina, USA
[15]Emory University School of Medicine, Atlanta, Georgia, USA
[16]University Medical Centre Utrecht, Wilhelmina Children's Hospital, Utrecht, The Netherlands
[17]Hannover Medical School, Hannover, Germany
[18]Universidade Estadual Paulista Júlio de Mesquita Filho, Botucatu, Saõ Paulo, Brazil
[19]Semmelweis University, Budapest, Hungary
[20]Royal Hospital for Sick Children, Glasgow, UK
[21]Royal Hospital for Sick Children, Edinburgh, UK
[22]Karolinska University Hospital, Stockholm, Sweden
[23]Hospital de Pediatría Garrahan, Buenos Aires, Argentina
[24]Istituto Giannina Gaslini, Pediatria II - Reumatologia, PRINTO, Genoa, Italy
[25]Charles University, Prague, Czech Republic

Acknowledgements We thank the following individuals for providing invaluable input and feedback on project development and support: Dr Daniel Aletaha (European League Against Rheumatism), Drs Suzette Peng and Sarah Yim (Food and Drug Administration), Drs Thorsten Vetter and Richard Vesely (European Medicines Agency), Bob Goldberg and Theresa Curry (The Myositis Association), Rhonda McKeever and Patti Lawler (Cure JM Foundation), and Irene Oakley (Myositis UK). We also thank Drs Michael Ward and Steven Pavletic for their critical review of the manuscript. Paul Hansen, who with Franz Ombler owns and co-invented the 1000Minds software referred to in the article, provided intellectual and logistic support for this project.

Collaborators Appendix A MEMBERS OF THE INTERNATIONAL MYOSITIS ASSESSMENT AND CLINICAL STUDIES GROUP AND THE PAEDIATRIC RHEUMATOLOGY INTERNATIONAL TRIALS ORGANISATION WHO CONTRIBUTED TO DEVELOPING THE RESPONSE CRITERIA

Steering committee: Lisa G Rider (co-principal investigator), Nicolino Ruperto (co-principal investigator), Rohit Aggarwal (methodology lead), Frederick W Miller, Jiri Vencovsky.

Statistical team: Rohit Aggarwal, Brian Erman, Nastaran Bayat, Angela Pistorio, Adam M. Huber, Brian M Feldman, Paul Hansen, Howard Rockette, Peter A Lachenbruch, Nicolino Ruperto, Lisa G. Rider.

Paediatric core set survey group: Maria Apaz, Suzanne Bowyer, Rolando Cimaz, Tamás Constantin, Megan Curran, Joyce Davidson, Brian M. Feldman, Thomas Griffin, Adam H. Huber, Olcay Jones, Susan Kim, Bianca Lang, Carol Lindsley, Daniel Lovell, Claudia Saad Magalhaes, Lauren M. Pachman, Clarissa Pilkington, Andrea Ponyi, Marilynn Punaro, Pierre Quartier, Athimalaipet V Ramanan, Angelo Ravelli, Ann Reed, Robert Rennebohm, David D Sherry, Clovis A Silva, Elizabeth Stringer, Annet van Royen-Kerkhof, Carol Wallace.

Clinical trial or natural history study data set contributors: Frederick W. Miller, Chester V. Oddis, Ann Reed, Lisa G. Rider, Nicolino Ruperto, and PRINTO members.

Paediatric patient profile working group: Maria Apaz, Tadej Avcin, Mara Becker, Michael W. Beresford, Rolando Cimaz, Tamás Constantin, Megan Curran, Ruben Cuttica, Joyce Davidson, Frank Dressler, Jeffrey Dvergsten, Sheila Knupp Feitosa de Oliveira, Brian M. Feldman, Virginia Paes Leme Ferriani, Berit Flato, Valeria Gerloni, Thomas Griffin, Michael Henrickson, Claas Hinze, Mark Hoeltzel, Adam M. Huber, Maria Ibarra, Norman Ilowite, Lisa Imundo, Olcay Jones, Susan Kim, Daniel Kingsbury, Bianca Lang, Carol Lindsley, Daniel Lovell, Alberto Martini, Claudia Saad Magalhaes, Bo Magnusson, Sheilagh Maguiness, Susan Maillard, Pernille Mathiesen, Liza McCann, Susan Nielsen, Lauren M. Pachman, Murray Passo, Clarissa Pilkington, Marilynn Punaro, Pierre Quartier, Egla Rabinovich, Athimalaipet V. Ramanan, Angelo Ravelli, Ann Reed, Robert Rennebohm, Lisa G. Rider, Rafael Rivas-Chacon, Angela Byun Robinson, Kelly Rouster-Stevens, Ricardo Russo, Lidia Rutkowska-Sak, Adriana Sallum, Helga Sanner, Heinrike Schmeling, Duygu Selcen, Bracha Shaham, David D. Sherry, Clovis A. Silva, Charles H. Spencer, Robert Sundel, Marc Tardieu, Akaluck Thatayatikom, Janjaap van der Net, Annet van Royen-Kerkhof, Dawn Wahezi, Carol Wallace, Francesco Zulian.

Conjoint analysis, paediatric group: Rolando Cimaz, Tamás Constantin, Ruben Cuttica, Joyce Davidson, Frank Dressler, Brian M. Feldman, Thomas Griffin, Michael Henrickson, Adam M. Huber, Lisa Imundo, Bianca Lang, Carol Lindsley, Claudia Saad Magalhaes, Bo Magnusson, Susan Maillard, Lauren M. Pachman, Murray Passo, Clarissa Pilkington, Marilynn Punaro, Angelo Ravelli, Ann Reed, Lisa G. Rider, Kelly Rouster-Stevens, Ricardo Russo, Bracha Shaham, Robert Sundel, Janjaap van der Net, Annet van Royen-Kerkhof.

Participants in consensus conference, paediatric working group: Rolando Cimaz, Rubén Cuttica, Sheila Knupp Feitosa de Oliveira, Brian M. Feldman, Adam M. Huber, Carol B. Lindsley, Clarissa Pilkington, Marilynn Punaro, Angelo Ravelli, Ann Reed, Kelly Rouster-Stevens, Annet van Royen-Kerkhof.

Participants in consensus conference, adult working group: Anthony Amato, Hector Chinoy, Robert G Cooper, Maryam Dastmalchi, Marianne de Visser, David Fiorentino, David Isenberg, James Katz, Andrew Mammen, Chester V. Oddis, Jiri Vencovsky, Steven R Ytterberg.

See Appendix A for members of the International Myositis Assessment and Clinical Studies Group and the Paediatric Rheumatology International Trials Organisation who contributed to developing the response criteria.

Contributors 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. LGR and NR 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: LGR, RA, FWM, JV, NR. Acquisition of data: LGR, RA, AP, NB, BE, BMF, AMH, RC, RJC, SKO, CBL, CAP, MP, AR, AMR, KR, AR, FD, CS, TC, JED, BM, RR, LV, MR, HR, PAL, JV, NR. Analysis and interpretation of data: LGR, RA, CAP, NB, BE, BMF, AMH, HR, PAL, FWM, JV, NR.

Competing interests None declared.

Provenance and peer review Not commissioned; internally peer reviewed.

Supported in part by the American College of Rheumatology, the European League Against Rheumatism, the NIH (Intramural Research Programs of the National Institute of Environmental Health Sciences (NIEHS), National Center for Advancing Translational Sciences, and National Institute of Arthritis and Musculoskeletal and Skin Diseases), Istituto G. Gaslini and the Paediatric Rheumatology International Trials Organisation (PRINTO), Cure JM Foundation, Myositis UK, and the Myositis Association. Dr. Vencovsky's work was supported by the Ministry of Health, Czech Republic (Institute of Rheumatology project for conceptual development of a research organisation, 00023728).

REFERENCES

1 Rider LG, Werth VP, Huber AM, *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 (IBMFRS), 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.
2 Miller FW, Rider LG, Chung YL, *et al*. Proposed preliminary core set measures for disease outcome assessment in adult and juvenile idiopathic inflammatory myopathies. *Rheumatology (Oxford)* 2001;40:1262–73.
3 Rider LG, Giannini EH, Harris-Love M, *et al*. Defining clinical improvement in adult and juvenile myositis. *J Rheumatol* 2003;30:603–17.
4 Ruperto N, Ravelli A, Murray KJ, *et al*. Preliminary core sets of measures for disease activity and damage assessment in juvenile systemic lupus erythematosus and juvenile dermatomyositis. *Rheumatology (Oxford)* 2003;42:1452–9.
5 Ruperto N, Ravelli A, Pistorio A, *et al*. The provisional Paediatric Rheumatology International Trials Organisation/American College of Rheumatology/European League Against Rheumatism Disease Activity Core Set for the Evaluation of Response to Therapy in Juvenile Dermatomyositis: a prospective validation study. *Arthritis Rheum* 2008;59:4–13.
6 Ruperto N, Martini A. Networking in paediatrics: the example of the Paediatric Rheumatology International Trials Organisation (PRINTO). *Arch Dis Child* 2011;96:596–601.
7 Singh G, Athreya BH, Fries JF, *et al*. Measurement of health status in children with juvenile rheumatoid arthritis. *Arthritis Rheum* 1994;37:1761–9.
8 Bode RK, Klein-Gitelman MS, Miller ML, *et al*. Disease Activity Score for children with juvenile dermatomyositis: reliability and validity evidence. *Arthritis Rheum* 2003;49:7–15.
9 Rider LG, Giannini EH, Brunner HI, *et al*. International consensus on preliminary definitions of improvement in adult and juvenile myositis. *Arthritis Rheum* 2004;50:2281–90.
10 Ruperto N, Pistorio A, Ravelli A, *et al*. The Paediatric Rheumatology International Trials Organization provisional criteria for the evaluation of response to therapy in juvenile dermatomyositis. *Arthritis Care Res (Hoboken)* 2010;62:1533–41.
11 Hasija R, Pistorio A, Ravelli A, *et al*. Therapeutic approaches in the treatment of juvenile dermatomyositis in patients with recent-onset disease and in those experiencing disease flare: an international multicenter PRINTO study. *Arthritis Rheum* 2011;63:3142–52.
12 Ruperto N, Pistorio A, Ravelli A, *et al*. Criteria to define response to therapy in paediatric rheumatic diseases. *Eur J Clin Pharmacol* 2011;67(Suppl 1):125–31.
13 Ruperto N, Pistorio A, Oliveira S, *et al*. Prednisone versus prednisone plus ciclosporin versus prednisone plus methotrexate in new-onset juvenile dermatomyositis: a randomised trial. *Lancet* 2016;387:671–8.
14 Felson DT, Furst DE, Boers M. Rationale and strategies for reevaluating the ACR20. *J Rheumatol* 2007;34:1184–7.
15 Rider LG, Ruperto N, Pistorio A, *et al*. 2016 development of adult dermatomyositis and polymyositis and juvenile dermatomyositis response criteria: methodological aspects: an American College of Rheumatology/European League Against Rheumatism/International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation Collaborative Initiative. *Rheumatology (Oxford)* 2016. In press.
16 Rider LG, Aggarwal R, Bayat N, *et al*. A hybrid conjoint analysis model is proposed as the definition of minimal, moderate and major clinical improvement in juvenile dermatomyositis clinical trials. *J Rheumatol* 2014;66(Suppl):S578.
17 Aggarwal R, Rider LG, Ruperto N, *et al*. A consensus hybrid definition using a conjoint analysis is the proposed response criteria for minimal and moderate improvement for adult polymyositis and dermatomyositis clinical trials. *Arthritis Rheumatol* 2014;66(Suppl):S404.
18 Aggarwal R, Rider LG, Ruperto N, *et al*. 2016 American College of Rheumatology/European League AGainst Rheumatism criteria for minimal, moderate, and major clinical response in adult dermatomyositis and polymyositis: International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation collaborative initiative. *Arthritis Rheumatol*. Published Online First.
19 Neogi T, Jansen TL, Dalbeth N, *et al*. 2015 gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheumatol* 2015;67:2557–68.
20 Utz KS, Hoog J, Wentrup A, *et al*. Patient preferences for disease-modifying drugs in multiple sclerosis therapy: a choice-based conjoint analysis. *Ther Adv Neurol Disord* 2014;7:263–75.
21 Amaya-Amaya M, Gerard K, Ryan M. Discrete choice experiments in a nutshell. In: Ryan M, Gerard K, Amaya-Amaya M, eds. *Using discrete choice experiments to value health and health care*. Dordrecht: Springer; 2008.
22 De Bekker-Grob E, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ* 2012;21:145–72.
23 De Lautour H, Taylor WJ, Adebajo A, *et al*. Development of preliminary remission criteria for gout using Delphi and 1000Minds consensus exercises. *Arthritis Care Res (Hoboken)* 2016;68:667–72.
24 Hansen P, Ombler F. A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *J Multi-Crit Decis Anal* 2008;15:87–107.
25 Oddis CV, Reed AM, Aggarwal R, *et al*. Rituximab in the treatment of refractory adult and juvenile dermatomyositis and adult polymyositis: a randomized, placebo-phase trial. *Arthritis Rheum* 2013;65:314–24.
26 Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283–98.
27 Finney DJ. *Statistical method in biological assay*. 3rd edn. London: Charles Griffin, 1978.
28 Liengme BV. *A guide to Microsoft Excel 2002 for scientists and engineers*. 3rd edn. Oxford: Butterworth-Heinemann, 2002.
29 Ruperto N, Meiorin S, Iusan SM, *et al*. Consensus procedures and their role in pediatric rheumatology. *Curr Rheumatol Rep* 2008;10:142–6.
30 Delbecq A, van de Ven A, Gustafson D. Group techniques for program planning: a guide to nominal group and Delphi processes. Glenview, IL: Scott, Foresman and Company, 1975.
31 Ruperto N, Ozen S, Pistorio A, *et al*. EULAR/PRINTO/PRES criteria for Henoch-Schonlein purpura, childhood polyarteritis nodosa, childhood Wegener granulomatosis and childhood Takayasu arteritis: Ankara 2008. Part I: overall methodology and clinical characterisation. *Ann Rheum Dis* 2010;69:790–7.
32 Piram M, Kone-Paut I, Lachmann HJ, *et al*. Validation of the auto-inflammatory diseases activity index (AIDAI) for hereditary recurrent fever syndromes. *Ann Rheum Dis* 2014;73:2168–73.
33 American College of Rheumatology Committee to Reevaluate Improvement Criteria. A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. *Arthritis Rheum* 2007;57:193–202.
34 Streiner DL. Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *Can J Psychiatry* 2002;47:262–6.
35 Felson DT, Anderson JJ, Boers M, *et al*. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38:727–35.
36 Giannini EH, Ruperto N, Ravelli A, *et al*. Preliminary definition of improvement in juvenile arthritis. *Arthritis Rheum* 1997;40:1202–9.
37 Consolaro A, Ruperto N, Bazso A, *et al*. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. *Arthritis Rheum* 2009;61:658–66.
38 Volochayev R, Csako G, Wesley R, *et al*. Laboratory test abnormalities are common in polymyositis and dermatomyositis and differ among clinical and demographic groups. *Open Rheumatol J* 2012;6:54–63.

2016 American College of Rheumatology/European League Against Rheumatism criteria for minimal, moderate, and major clinical response in adult dermatomyositis and polymyositis

An International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation Collaborative Initiative

Rohit Aggarwal,[1] Lisa G Rider,[2] Nicolino Ruperto,[3] Nastaran Bayat,[2] Brian Erman,[4] Brian M Feldman,[5] Chester V Oddis,[1] Anthony A Amato,[6] Hector Chinoy,[7] Robert G Cooper,[8] Maryam Dastmalchi,[9] David Fiorentino,[10] David Isenberg,[11] James D Katz,[2] Andrew Mammen,[12] Marianne de Visser,[13] Steven R Ytterberg,[14] Ingrid E Lundberg,[9] Lorinda Chung,[10] Katalin Danko,[15] Ignacio García-De la Torre,[16] Yeong Wook Song,[17] Luca Villa,[3] Mariangela Rinaldi,[3] Howard Rockette,[1] Peter A Lachenbruch,[2] Frederick W Miller,[2] Jiri Vencovsky,[18] for the International Myositis Assessment and Clinical Studies Group and the Paediatric Rheumatology International Trials Organisation

Handling editor Tore K Kvien

For numbered affiliations see end of article.

Correspondence to
Dr Rohit Aggarwal, Division of Rheumatology and Clinical Immunology, Department of Medicine, University of Pittsburgh, 3601 5th Avenue, Suite 2B, Pittsburgh, PA 15261, USA; aggarwalr@upmc.edu

See Appendix A for members ofthe International Myositis Assessment and Clinical Studies Group and the Paediatric Rheumatology International Trials Organisation who contriuted to developing the response criteria.

RA and LGR contributed equally and FWM and JV contributed equally.

This article is published simultaneously in the May 2017 issue of *Arthritis & Rheumatology*.

Submitted for publication 11 February 2016; accepted in revised form 31 January 2017

Accepted 1 March 2017
Published Online First 31 March 2017

ABSTRACT

To develop response criteria for adult dermatomyositis (DM) and polymyositis (PM). Expert surveys, logistic regression, and conjoint analysis were used to develop 287 definitions using core set measures. Myositis experts rated greater improvement among multiple pairwise scenarios in conjoint analysis surveys, where different levels of improvement in 2 core set measures were presented. The PAPRIKA (Potentially All Pairwise Rankings of All Possible Alternatives) method determined the relative weights of core set measures and conjoint analysis definitions. The performance characteristics of the definitions were evaluated on patient profiles using expert consensus (gold standard) and were validated using data from a clinical trial. The nominal group technique was used to reach consensus. Consensus was reached for a conjoint analysis-based continuous model using absolute per cent change in core set measures (physician, patient, and extramuscular global activity, muscle strength, Health Assessment Questionnaire, and muscle enzyme levels). A total improvement score (range 0–100), determined by summing scores for each core set measure, was based on improvement in and relative weight of each core set measure. Thresholds for minimal, moderate, and major improvement were ≥20, ≥40, and ≥60 points in the total improvement score. The same criteria were chosen for juvenile DM, with different improvement thresholds. Sensitivity and specificity in DM/PM patient cohorts were 85% and 92%, 90% and 96%, and 92% and 98% for minimal, moderate, and major improvement, respectively. Definitions were validated in the clinical trial analysis for differentiating the physician rating of improvement ($p<0.001$). The response criteria for adult DM/PM consisted of the conjoint analysis model based on absolute per cent change in 6 core set measures, with thresholds for minimal, moderate, and major improvement.

Idiopathic inflammatory myopathies are a group of acquired, heterogeneous, systemic connective tissue diseases that include adult dermatomyositis (DM) and polymyositis (PM) and juvenile DM.[1] Despite significant morbidity and mortality associated with DM/PM, there are currently no therapies approved for these syndromes by the Food and Drug Administration or the European Medicines Agency based on randomised controlled trials. However, with the advancement in novel therapeutic agents that target various biologic pathways implicated in the pathogenesis of DM/PM,[2] there is a need for

To cite: Aggarwal R, Rider LG, Ruperto N, *et al*. *Ann Rheum Dis* 2017;**76**:792–801.

This criteria set has been approved by the American College of Rheumatology (ACR) Board of Directors and the European League Against Rheumatism (EULAR) Executive Committee. This signifies that the criteria set has been quantitatively validated using patient data, and it has undergone validation based on an independent data set. All ACR/EULAR-approved criteria sets are expected to undergo intermittent updates.

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

well-designed clinical trials using validated and universally accepted outcome measures. Recently completed clinical trials in adult DM/PM and juvenile DM have used varying response criteria,[3–5] again highlighting the need for both data- and consensus-driven criteria to be used uniformly in future studies. Core set measures of myositis disease activity for adult DM/PM clinical trials have been established and validated by the International Myositis Assessment and Clinical Studies Group (IMACS);[6–8] these measures were used as the foundation for the current study. We undertook this study because there is a need for composite response criteria in myositis, given the heterogeneity of the disease and the fact that no single core set measure adequately covers all the domains in myositis. For example, muscle enzyme levels can be normal in active DM, and active muscle weakness in DM can occur without active rash.

Preliminary response criteria for adult DM/PM had been developed and partially validated by IMACS; these criteria were based on at least 20% improvement in 3 of 6 core set measures, with no more than 2 core set measures worsening by at least 25% (which cannot be muscle strength).[8 9] However, those criteria were considered preliminary, because they were not prospectively validated. Moreover, newer methodologies such as conjoint analysis and other continuous or hybrid approaches for developing response criteria had not been evaluated.[10–14] The preliminary criteria had other potential limitations, including equal weights being applied to each core set measure and the lack of quantitative or continuous outcomes. With the growing repertoire of potential therapeutic agents, some of which may yield better results than only minimal clinical improvement, there is also a need to develop criteria for moderate and major clinical improvement.

For these reasons, and with support from the American College of Rheumatology, European League Against Rheumatism, IMACS, and the Paediatric Rheumatology International Trials Organisation (PRINTO),[15] a collaboration was established to develop a data- and consensus-driven process involving multiple clinical data sets and the international myositis community in order to develop and validate response criteria for adult DM/PM and juvenile DM. This effort involved a comprehensive approach to developing candidate definitions for the response criteria, including continuous or hybrid definitions, using conjoint analysis,[13 14 16–19] and for developing criteria for minimal as well as greater degrees of improvement. This article focuses on the criteria for minimal and moderate improvement for adult DM/PM, whereas the threshold for major improvement is considered preliminary. A companion article focuses on the juvenile DM response criteria.[20]

METHODS

Core set measures and patient profile consensus

To develop patient profiles as well as candidate definitions for response criteria in adult PM and DM, we used previously validated IMACS myositis core set measures for patients with adult DM/PM, which include physician and patient global activity on a 10-cm visual analogue scale (VAS), muscle strength measured by manual muscle testing (MMT), physical function measured by the Health Assessment Questionnaire (HAQ),[21] extramuscular global activity measured by the physician on a 10-cm VAS, and the most abnormal serum muscle enzyme.[8 22] The entire process, from the development of these profiles and candidate definitions through final consensus voting, is shown in the flow diagram in figure 1.[23 24] Details of the methodology used to develop patient profiles, candidate definitions, validation, and expert consensus will be described in a separate publication.[24] Briefly, patient data from natural history studies and uncontrolled clinical trials were used to develop patient profiles, which were then rated by adult myositis experts to achieve consensus as to whether improvement was none, minimal, moderate, or major. The expert consensus of improvement was used as the gold standard to validate various

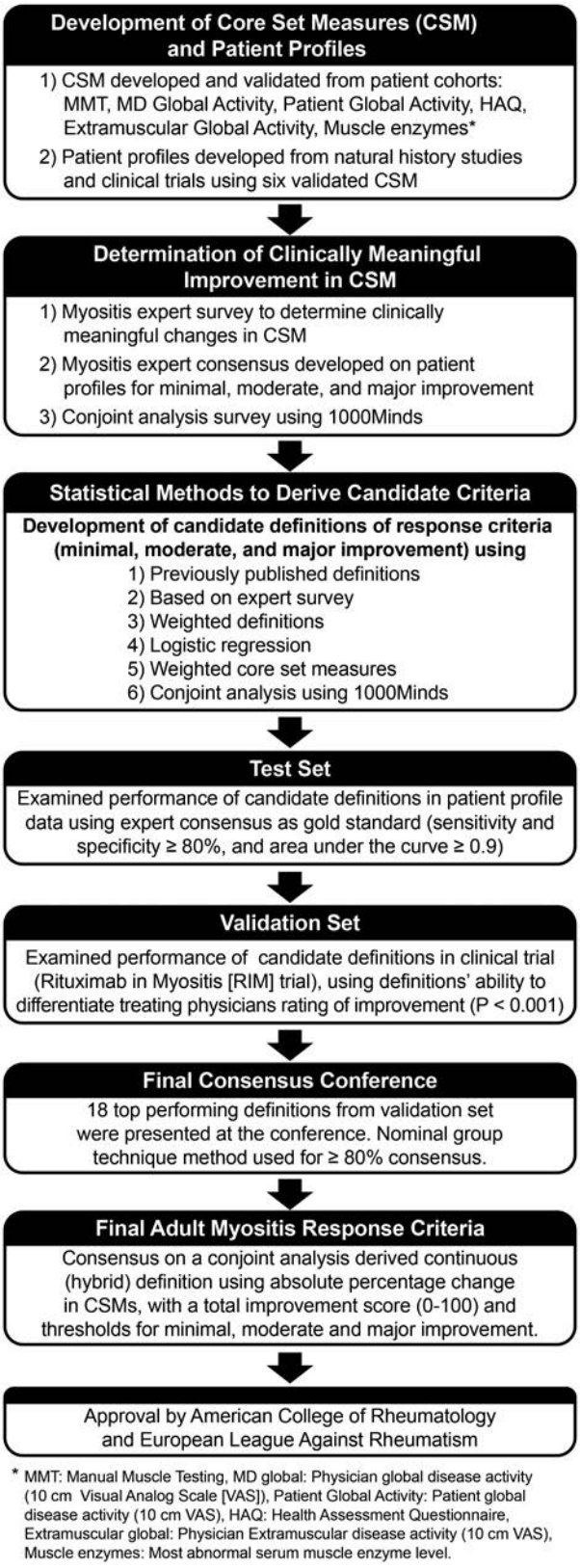


Figure 1 Flow diagram of the entire process used to develop and validate the approved response criteria for adult dermatomyositis and polymyositis.

candidate definitions. The Bohan and Peter classification as used to designate definite or probable adult DM/PM.[25]

Candidate definitions of response criteria

Six different types of candidate definitions for minimal, moderate, and major response (table 1) were developed:[23 26] 3 types of definitions were traditional (categorical), and 3 were continuous (hybrid). Traditional definitions provide only categorical outcomes of minimal, moderate, and major improvement, or not improved, based on the criteria, whereas continuous definitions yield an improvement score as a continuous outcome measure, with thresholds of minimal, moderate, and major improvement serving as categorical outcomes. Continuous definitions are considered hybrid definitions, because the same definition can be used as a continuous or categorical outcome measure based on the study requirements. Definitions utilising either absolute per cent change (final minus baseline divided by range and multiplied by 100) or relative per cent change (final minus baseline, divided by baseline and multiplied by 100) were evaluated as candidate definitions.

Conjoint analysis surveys

Conjoint analysis surveys were administered to myositis experts using 1000Minds online software.[11] Experts were presented with pairs of hypothetical patient scenarios; each patient had different levels of improvement in the same 2 core set measures, assuming other core set measures remained the same. Experts rated which of the 2 scenarios had greater improvement. Based on the rater's response, all other hypothetical patients that could be pairwise ranked were eliminated via the property of transitivity, thereby significantly reducing the number of scenarios presented. The PAPRIKA (Potentially All Pairwise Rankings of All

Table 1 Types of candidate definitions for response criteria that were developed and tested

Type of candidate definitions of response	Description	Example of candidate definition for the response criteria
Previously published (categorical definition)	Previously published definitions of improvement that were retested	Minimal. Three of any 6 improved by ≥20%, no more than 2 worse by >25% (which cannot be MMT)[9] Moderate. Three of any 6 improved by ≥50%, no more than 2 worse by >25% (which cannot be MMT) Major. Three of any 6 improved by ≥70%, no more than 2 worse by >25% (which cannot be MMT)
Newly drafted (categorical definition)	Drafted relative or absolute % change candidate definitions of response, based on recent CSM survey	Minimal. Two of any 6 improved by ≥30%, no more than 1 worse by >30% (which cannot be MMT) Moderate. Two of any 6 improved by ≥50%, no more than 1 worse by >30% (which cannot be MMT) Major. Two of any 6 improved by ≥75%, no more than 1 worse by >30% (which cannot be MMT)
Weighted (categorical definition)	Applied conjoint analysis relative weights to CSM in newly drafted definitions; each CSM receives improvement points (corresponding relative weights), when it reaches the threshold for minimal, moderate, or major improvement; worsening points are applied similarly; improvement is calculated based on a total score of improvement versus worsening	Improvement=at least 2.5 total improvement points of a maximum possible score of 8, and no more than 2.5 worsening points, where MD global=1.5 points, patient global=1 point, MMT=2 points, HAQ=1.5 points, extramusc=1.5 points, enzyme=0.5 point Minimal. Improvement points given when CSM ≥30%; worsening points given when CSM worse by >25% Moderate. Improvement points given when CSM ≥50%; worsening points given when CSM worse by >25% Major. Improvement points given when CSM ≥75%; worsening points given when CSM worse by >25%
Logistic regression (continuous definition)	Model of improvement using combination of CSM with different weights, as developed in the logistic regression model and rounded for better feasibility; total scores derived, with different cutoffs, for minimal, moderate, and major improvement	Improvement score=5×(MD global % change)+3×(patient global % change)+(MMT % change)+2×(HAQ % change)+2×(extramusc % change)+2.5×(enzyme % change) Minimal. Improvement score ≥250 Moderate. Improvement score ≥500 Major. Improvement score ≥750
Core set measure–weighted (continuous definition)	Multiply the % change in each CSM by the weights derived from conjoint analysis, then sum (% change in each CSM×conjoint analysis weights) to get final total improvement score; different thresholds for minimal, moderate, and major improvement established based on consensus profile ratings as gold standard	Improvement score=2×(MD global % change)+(patient global % change)+3×(MMT % change)+1.5×(HAQ % change)+1.5×(extramusc % change)+(enzyme % change) Minimal. Improvement score ≥100 Moderate. Improvement score ≥250 Major. Improvement score ≥400
Conjoint analysis (continuous definition)	For a given range in the level of improvement in each CSM, a score is assigned, as developed by the conjoint-analysis survey results and modelling; greater degrees of improvement receive higher scores; a patient is minimally improved if the improvement score is above the cutoff for minimal improvement; similarly, for moderate and major improvement	Cut points for the model are: Minimal. Improvement score ≥20 Moderate. Improvement score ≥40 Major. Improvement score ≥60*

*See table 3 for cut points for the full model.

CMAS, Childhood Myositis Assessment Scale; CSM, core set measure; enzyme, most abnormal serum muscle enzyme value among aldolase, alanine aminotransferase, aspartate aminotransferase, lactate dehydrogenase, and creatine kinase; extramusc, extramuscular global activity; HAQ, Health Assessment Questionnaire; MD global, physician global activity score; MMT, manual muscle testing; patient global, patient global activity score.

Possible Alternatives) method was used to determine the relative importance of the core set measures. Relative weights of core set measures and their levels of improvement were used to develop a scoring system by mathematical methods based on linear programming,[13] such that when all 6 core set measures are considered together, the maximum score (total improvement score) possible for representing a patient's improvement is 100 and the minimum score is 0. The thresholds for minimal, moderate, and major improvement in the total improvement score were based on optimum sensitivity and specificity (using the Youden index[27]) in the subset of patient cohort data.

Validation of candidate response criteria

The performance characteristics of candidate criteria were evaluated using consensus profile ratings as the gold standard, assessing sensitivity, specificity, and area under the curve (AUC) to compare the performance of these candidate definitions. Those that performed well in the consensus profiles (sensitivity and specificity ≥80%, AUC ≥0.9 for minimal improvement, and AUC ≥0.8 for moderate and major improvement) were externally validated using data for adult DM/PM patients (n=142) enrolled in the Rituximab in Myositis (RIM) trial.[3] The treating physician's rating of improvement (0–7 scale) at 24 weeks in the RIM trial was used for validation, and a 1-point change in the physician's rating was considered clinically significant.[3] We then selected the top candidate definitions (up to 4 top-performing definitions from each of the 6 different types of candidate definitions) for consideration at the final consensus conference, in order to discuss a manageable number of definitions at the conference.

Consensus conference

The nominal group technique (NGT) was applied to develop consensus among experts in adult DM/PM regarding the top-performing candidate definitions for minimal and moderate improvement in adult DM/PM.[28–30] Experienced moderators (RA and FWM) led the NGT consensus-development process for the adult working group and the combined adult and paediatric working group (RA, LGR, NR, and FWM). Given the paucity of data on major improvement, we considered the major improvement thresholds as preliminary for the final consensus meeting. For each candidate definition, the methodologic details used to develop them it and its performance characteristics in the consensus patient profiles and the RIM trial were presented to the adult working group. Each of the 12 participants in the adult working group independently reviewed the performance characteristics of all 18 top candidate definitions for adult DM/PM. Detailed data for each candidate definition, including sensitivity, specificity, and AUC as well as kappa values and ORs for minimal, moderate, and major improvement, were provided. The AUC was determined from the receiver operating characteristic curve as a plot of sensitivity versus (1—specificity) for total improvement scores as well as for thresholds.[27]

Adult working group

The primary goal for the adult working group was to develop consensus response criteria for minimal and moderate clinical improvement in adult DM/PM based on the data presented, as well as the face validity, feasibility, and generalisability of the proposed candidate criteria. The experts in the adult working group included internationally recognised rheumatologists, neurologists, and dermatologists who have considerable experience in myositis and with the core set measures. Voting was conducted in an independent, anonymous, and systematic manner via a web-based system developed by staff at the PRINTO coordinating centre.[31 32] In the initial rounds of voting, participants were asked to rank their top 5 choices. The results were compiled, and aggregate votes and rank of each candidate definition were shared with the group after each round of voting. Participants were then asked in a random manner to discuss their top-ranked and bottom-ranked choices. Candidate definitions receiving a small proportion of votes were eliminated. In subsequent voting rounds, participants were asked to re-rank their choices after reviewing the previous round's voting and discussion. When fewer than 5 candidate definitions remained, each participant selected one as the top response criteria. The objective was to continue the rounds of voting in the same manner until a single candidate definition reached consensus (≥80% of the votes) or until it was clear that consensus would not be reached.

Combined adult and paediatric working group

After consensus was achieved by each working group, both groups then came together to vote on common response criteria to be used for both adult DM/PM and juvenile DM[20] as the outcome measure for combined clinical trials. For this voting round, the top candidate definitions from the final round of voting in each working group were considered, and a similar online voting system and the NGT were used until consensus of ≥80% was reached (28–30). For determining the thresholds of improvement for the selected definition, the required consensus was ≥70%, which was done by post-conference voting.

RESULTS

Candidate definitions

A total of 287 adult DM/PM candidate response criteria were drafted or derived using data-driven methods. Included were 10 previously published definitions, 134 newly drafted definitions based on expert survey results, 63 weighted definitions, 68 logistic regression definitions, 6 conjoint analysis definitions, and 6 definitions in which differential weights were applied to the improvement achieved in each core set measure. Among these definitions, 163 used relative per cent change and 124 used absolute per cent change in the core set measures.

Validation

Candidate definitions with a sensitivity and specificity of ≥80%, AUC ≥0.9 for minimal, and AUC ≥0.8 for moderate and major improvement in the patient profile analysis using expert consensus rating as the gold standard were evaluated for external validation using RIM clinical trial data[3] (see online supplementary table S1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40064/abstract). Thus, of 122 adult DM/PM candidate definitions evaluated using the RIM trial data, 36 adult DM/PM candidate definitions, including 25 using relative and 11 using absolute per cent change in core set measures, had AUC ≥0.7 and showed validation in the clinical trial analysis.

Top candidate definitions

Of 36 validated definitions, 17 top-performing adult candidate definitions and the top paediatric response criteria[20] were considered by the adult working group at the consensus conference so that, in total, 18 candidate definitions were evaluated (table 2 and see online supplementary table S2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40064/abstract). They included 9 categorical definitions and 9 continuous definitions, in which 14 used relative

Table 2 Detailed performance characteristics of patient profiles and clinical trial data for the top 5 candidate response criteria definitions presented at the consensus conference*

	Profiles (n=270)†				RIM trial (n=147)			
Candidate definitions for response criteria, improvement category, core set measure	Sensitivity, %	Specificity, %	Threshold AUC	Total AUC	Candidate definition, improved physician's rating‡	Candidate definition, not improved physician's rating‡	p Value	Rank
Conjoint analysis absolute % change (model 3)§								1
Minimal (improvement score ≥20)	85	92	0.89	0.96	2.0	4.0	<0.001	
Moderate (improvement score ≥40)	90	96	0.93	0.99	2.0	3.0	<0.001	
Major (total improvement score ≥60)	92	98	0.95	1.00	2.0	3.0	<0.001	
Conjoint analysis relative % change (model 1)¶								2
Minimal (improvement score ≥33)	94	90	0.92	0.98	2.0	4.0	<0.001	
Moderate (improvement score ≥55)	93	93	0.93	0.99	2.0	3.0	<0.001	
Major (improvement score ≥70)	100	95	0.97	0.99	2.0	3.0	<0.001	
Conjoint analysis relative % change (model 2)¶								3
Minimal (improvement score ≥30)	94	92	0.93	0.98	2.0	4.0	<0.001	
Moderate (total improvement score ≥45)	94	88	0.91	0.98	2.0	3.0	<0.001	
Major (improvement score ≥65)	100	98	0.99	1.00	2.0	3.0	<0.001	
Weighted core set measure relative % change**								4
Minimal (improvement score ≥100)	92	91	0.91	0.97	2.0	3.0	<0.001	
Moderate (improvement score ≥250)	94	91	0.93	0.98	2.0	3.0	<0.001	
Major (improvement score ≥400)	100	94	0.97	1.00	2.0	3.0	<0.001	
Logistic regression relative % change††								5
Minimal (improvement score ≥75)	89	93	0.91	0.97	2.0	3.0	<0.001	
Moderate (improvement score ≥150)	94	88	0.91	0.98	2.0	3.0	<0.001	
Major (improvement score ≥300)	100	96	0.98	1.00	2.0	3.0	<0.001	

*Online supplementary table S2 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40064/abstract) shows definitions 6–18 from the consensus conference ratings. The threshold AUC was calculated as the AUC from the ROC curve for the total improvement score and the threshold for minimal, moderate, and major improvement. The total AUC was calculated as the AUC from the ROC curve, using the total improvement score and the threshold cutoffs for minimal, moderate, and major improvement, and applies only to continuous definitions.
†The reference standard for sensitivity and specificity was myositis expert consensus rating of improvement.
‡Physician's rating is the treating physician's rating on a Likert scale of 1–7, where lower scores represent a greater degree of improvement, at week 24 of the RIM trial.[3] A 1-point difference in the physician's rating of improvement from no improvement to minimal improvement was considered not only statistically significant but also clinically significant.
§Conjoint analysis–based continuous candidate response criteria using absolute per cent change in core set measures (absolute per cent change model) is shown in table 3. These criteria are also the top response criteria for juvenile DM, but with different thresholds in the total improvement score for minimal, moderate and major improvement.[20]
¶Conjoint analysis–based continuous candidate response criteria using relative per cent change in core set measures are shown in online supplementary table S3 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40064/abstract). These criteria are also the second- and third-choice criteria for juvenile DM, but with different thresholds in the total improvement score for minimal, moderate, and major improvement.[20]
**The total improvement score is calculated as 2×(MD global % change)+(patient global % change)+3×(MMT % change)+1.5×(HAQ % change)+1.5×(extramusc % change)+(enzyme % change).
††The total improvement score is calculated as (MD global % change)+(patient global % change)+(MMT % change)+(HAQ % change)+(extramusc % change)+enzyme % change).
AUC, area under the curve; DM, dermatomyositis; extramusc, extramuscular; enzyme, most abnormal serum muscle enzyme value among aldolase, alanine aminotransferase, aspartate aminotransferase, lactate hydrogenase, and creatine kinase; HAQ, Health Assessment Questionnaire; MD global, physician global activity; MMT, manual muscle testing; patient global, patient global activity; RIM, Rituximab in Myositis; ROC, receiver operating characteristic.

per cent change and 4 used absolute per cent change in core set measures. In each categorical definition, a patient would either meet or not meet the response criteria of minimal, moderate, or major improvement based on the degree of improvement or worsening in each core set measure. In the continuous definitions, however, each subject generates a total improvement score on a continuous scale, such that a greater degree of improvement corresponds to a higher score. Furthermore, patients could be categorised as achieving minimal, moderate, or major clinical improvement based on reaching the pre-set threshold score on the continuous scale. Table 2 shows the performance characteristics of the top 5 candidate definitions for the response criteria selected at the consensus conference (see online supplementary table S2 for definitions 6–18).

In the patient profiles, with expert consensus as the gold standard, all top candidate definitions presented at the conference had excellent performance characteristics, with median sensitivity of 87% (IQR 84–90%) and specificity of 94% (IQR 92–95%) for minimal improvement with a median AUC of 0.91 (IQR 0.90–0.92) (table 2 and see online supplementary tables S1 and S2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40064/abstract).

Sensitivity, specificity, and AUC were similarly high for moderate and major improvement criteria for these definitions (table 2 and see online supplementary tables S1 and S2). All candidate definitions presented at the conference were validated using the RIM trial data at the 24-week time point and were shown to differentiate ($p<0.001$) between the treating physician's improvement score at week 24 in patients rated as improved versus not improved[3] (table 2 and see online supplementary tables S1 and S2).

Consensus conference voting

The top-choice definition for the adult working group, which received 80% of the votes, was the conjoint analysis-based continuous definition model 1, which includes relative per cent change in core set measures, including physician and patient global activity, muscle strength, physical function, most abnormal serum enzyme level, and extramuscular activity (see online supplementary table S3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.

1002/art.40064/abstract). The second-choice definition, receiving 20% of the votes, was the conjoint analysis-based continuous model 2, which also includes relative per cent change in core set measures (see online supplementary table S3). Models 1 and 2 differ only in the scores associated with each level of improvement in each core set measure.

However, in the final round of voting and discussion, adult working group participants reached unanimous consensus that the response criteria for adult DM/PM would be identical to the top-choice response criteria for juvenile DM, which is a conjoint analysis-based continuous definition (model 3) using absolute per cent change in core set measures (table 3).[20] Participants favoured using the same response criteria for adult DM/PM and juvenile DM so that data from different studies can be harmonised more effectively and to facilitate combined trials, especially given that the definitions were similar with similar performance characteristics. Moreover, the absolute per cent change in core set measures (model 3, table 3) was thought to be more representative of meaningful clinical change compared with relative per cent change in core set measures (models 1 and 2, supplementary table 3). Participants also voted to evaluate all top 5 candidate definitions from the adult working group in future clinical trials, with the other 4 as secondary outcome measures. The top 3 of these criteria, the conjoint analysis definitions, are the same for both adult DM/PM and juvenile DM, with different thresholds of improvement.

The sensitivity and specificity of the top-choice criteria, the conjoint analysis absolute percent change (table 3), were 85% and 92% for minimal improvement, 90% and 96% for moderate improvement, and 92% and 98% for major improvement, respectively (table 2). The AUC was 0.96 for the total improvement score and 0.89, 0.93, and 0.95 for minimal, moderate, and major improvement thresholds, respectively (table 2). In the RIM trial,[3] these response criteria showed a significant difference in the physician's rating of improvement when the response criteria rated the patient as improved versus not improved for minimal, moderate, and major improvement ($p<0.001$) (table 2 and see online supplementary table S2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.40064/abstract). Myositis experts in the consensus conference favoured the conjoint analysis-based continuous response criteria because the total improvement score is a continuous measure that corresponds to the magnitude of improvement in a patient and provides the ability to categorise a patient's degree of improvement as minimal, moderate, or major (making it truly a hybrid definition). Moreover, the differential weights for various core set measures were also thought to be congruent with an expert's assessment of the relative importance of each core set measure. An important consideration in the final selection was that the top-choice definition be based on absolute per cent change in the core set measure, which was favoured by the participants because, given the various VAS measurements used, the absolute per cent change was thought to be more representative of meaningful clinical change.

Top candidate definitions considered by the combined paediatric/adult working group

Three candidate definitions were considered by the combined adult/paediatric working group; these included the top adult definitions (see online supplementary table S3) and the top paediatric definitions,[20] one of which was identical in both groups. Final consensus was reached for the combined adult DM/PM and juvenile DM response criteria, with 91% of

Table 3 Final myositis response criteria for minimal, moderate, and major improvement in adult dermatomyositis/polymyositis (DM/PM) and combined adult DM/PM and juvenile DM clinical trials and studies*

Core set measure, level of improvement based on absolute percent change	Improvement score
Physician global activity	
Worsening to 5% improvement	0
>5–15% improvement	7.5
>15–25% improvement	15
>25–40% improvement	17.5
>40% improvement	20
Patient global activity	
Worsening to 5% improvement	0
>5–15% improvement	2.5
>15–25% improvement	5
>25–40% improvement	7.5
>40% improvement	10
Manual muscle testing	
Worsening to 2% improvement	0
>2–10% improvement	10
>10–20% improvement	20
>20–30% improvement	27.5
>30% improvement	32.5
Health Assessment Questionnaire	
Worsening to 5% improvement	0
>5–15% improvement	5
>15–25% improvement	7.5
>25–40% improvement	7.5
>40% improvement	10
Enzyme (most abnormal)	
Worsening to 5% improvement	0
>5–15% improvement	2.5
>15–25% improvement	5
>25–40% improvement	7.5
>40% improvement	7.5
Extramuscular activity	
Worsening to 5% improvement	0
>5–15% improvement	7.5
>15–25% improvement	12.5
>25–40% improvement	15
>40% improvement	20

The **total improvement score** is the sum of all 6 improvement scores associated with the change in each core set measure. A total improvement score of **≥20** represents **minimal improvement**, a score of **≥40** represents **moderate improvement**, and a score of **≥60** represents **major improvement**.

*Note that these response criteria are also proposed for use in combined adult DM/PM and juvenile DM trials (20). For comparison, the thresholds of improvement in the total improvement score for juvenile DM are ≥30 for minimal improvement, ≥45 for moderate improvement, and ≥70 for major improvement. Also note that the criteria for major improvement for adult DM/PM are preliminary.

How to calculate the improvement score: The absolute percent change [final value −baseline value/range×100) is calculated for each core set measure. For muscle enzymes, the most abnormal serum muscle enzyme value at baseline (creatine kinase, aldolase, alanine transaminase, aspartate aminotransferase, lactate dehydrogenase) is used. The enzyme range was calculated based on a 90% range of enzymes from natural history data,[34 46] which for creatine kinase is 15 times the upper limit of normal (ULN), for aldolase is 6 times the ULN, and for lactate dehydrogenase, aspartate aminotransferase, and alanine transaminase is 3 times the ULN. The ULN is determined according to the individual laboratories in the participating centres. The ranges for physician global activity, patient global activity, manual muscle testing, and extramuscular global activity are based on the instrument scale used.[3 26] An improvement score is assigned for each core set measure based on the absolute per cent change in the core set measure according to the definition. These individual core set measure improvement scores are then totalled among the 6 core set measures to give the total improvement score. The thresholds for minimal, moderate, and major improvement are provided. The total improvement score itself may also be compared among treatment arms in a trial. A total improvement score between 0 and 100 corresponds to the degree of improvement, with higher scores corresponding to a greater degree of improvement.

participants voting for the conjoint analysis-based continuous definition, based on absolute per cent change in the core set measure (table 3). The combined working group agreed that the same final response criteria will be used for clinical trials of both adult DM/PM and juvenile DM, but with different thresholds for improvement in adult versus paediatric patients as well as different core set measures for adult patients (IMACS) and paediatric patients (IMACS and PRINTO). Participants favoured using the same response criteria for adult DM/PM and juvenile DM, because the top definition from each working group was very similar (ie, both being conjoint analysis-based continuous models, with excellent and similar performance characteristics) and because it would permit comparison of outcomes in separate studies. Although only the IMACS core set measures were used for adult DM/PM, for further congruence with paediatric core set measures, the experts in adult myositis agreed to include the Short Form-36[33] as a health-related quality-of-life measure to correspond to the PRINTO quality-of-life core set measure, the parent form of the Child Health Questionnaire.[34–36] In a post-conference final vote, consensus (74%) was reached on threshold values for minimal, moderate, and major response for adult DM/PM patients, which are ≥20 in the total improvement score for minimal improvement, ≥40 for moderate improvement, and ≥60 for major improvement. In contrast, consensus on the final threshold values for minimal, moderate, and major response for juvenile DM were ≥30, ≥45, and ≥70 points, respectively.

DISCUSSION

After a systematic data- and consensus-driven process, a conjoint analysis-based continuous (ie, hybrid) definition based on absolute per cent change in core set measures was selected as the response criteria for adult DM/PM for minimal and moderate improvement in future clinical trials and studies (figure 1). Because the total number of cases in the trial data sets and clinical profiles that achieved major improvement was small, it was decided that the thresholds for major improvement would be considered preliminary. The same continuous (or hybrid) definition, but with different thresholds for minimal, moderate, and major improvement in IMACS or PRINTO core set measures, will be used for juvenile DM clinical trials and studies, as well as for combined adult DM/PM and juvenile DM studies and clinical trials in the future.[20 24]

The process for developing and validating the candidate definitions for the response criteria was extensive and comprehensive, as we used large prospective clinical cohort data sets to develop patient profiles, and myositis expert consensus was used as the gold standard for clinical response. Consequently, we derived six different types of candidate definitions, each with many variations, leading to a total of 287 candidate definitions tested, which were validated using natural history cohorts and data from a randomised clinical trial. Subsequently, a representative number of international myositis experts from various disciplines (rheumatology, neurology, and dermatology) agreed on an innovative continuous (or hybrid) model using absolute per cent change in validated core set measures.

These response criteria were developed using a novel conjoint analysis methodology, the 1000Minds software.[13] Conjoint analysis, or discrete choice experiment, is a statistical technique used to determine expert group decision-making around various measures (and multiple levels within each measure), providing the ability to develop differential weighting of measures and composite criteria using those measures. The 1000Minds software for conjoint analysis has been used recently to develop rheumatologic classification and/or outcome criteria for rheumatoid arthritis (RA), systemic sclerosis,[12 13 37 38] and gout.[11 16 17 39]

The criteria developed are continuous in nature and generate a total improvement score (on a scale of 0–100), which can provide a quantitative degree of improvement for each patient rather than a dichotomous or categorical assessment of improvement. The total improvement score is the sum of the improvement reflected in each of the 6 core set measures, but the individual core set measures are weighted, such that those deemed more important provide a greater contribution to the final score. For example, changes in the MMT and physician global disease activity scores are weighted more heavily than changes in the most abnormal enzyme or the HAQ. These weights were consistent with our myositis expert survey,[26] which was independent of the process used to develop and validate our response criteria.

There are significant advantages of using continuous response criteria (especially in pilot studies). For example, it might be possible to enrol fewer subjects and still have sufficient statistical power to differentiate between treatment groups by using the mean or median total improvement score. Moreover, continuous measures have the best sensitivity to change, the use of which allows modest treatment differences to be detected as statistically significant, which in turn leads to better clinical trials.[10] Moreover, the criteria developed provide thresholds for both minimal and moderate improvement, with a preliminary threshold for major improvement. Therefore, larger, adequately powered clinical trials and studies can use the threshold of minimal clinically significant improvement to differentiate the treatment groups, because this difference will be considered *clinically* significant. Similarly, the proportions of patients achieving minimal or moderate improvement can be determined and compared between treatment arms. The ability of the same response criteria to be used not only as a continuous measure, where a higher score implies greater improvement, but also as a categorical response of minimal and moderate improvement, results in a unique hybrid aspect to these criteria.

Another advantage of continuous response criteria over the previous IMACS response criteria is that inclusion criteria for clinical trials will not require minimal severity in any core set measure, because all levels of improvement in each core set measure contribute more or less to the response. However, for each trial the investigators will have to determine the entry criteria for baseline core set measure abnormality, but those will depend on the effect size, disease or organ target, recruitment, and feasibility rather than on the response criteria alone. This is an improvement over the previous IMACS preliminary response criteria, where the clinical trial inclusion criteria required a baseline deficit of at least 20% in each core set measure to enable reaching the threshold of ≥20% improvement in core set measures after treatment.

Another important aspect of these response criteria is that they are based on an absolute per cent change in core set measures rather than relative per cent change, as used for scoring other rheumatologic diseases such as RA[40 41] and prior myositis response criteria.[9] The panellists strongly believed that absolute per cent change rather than relative per cent change in core set measures more accurately reflects the degree of change. For example, for a patient in whom disease activity improved from 2 to 1 cm on a 10-cm VAS, this was interpreted by experts as more consistent with 10% improvement (absolute per cent change) and not as 50% improvement reflected by relative per cent change. Also, because many of the myositis core set

measures arbitrarily have 0 as the lower limit of normal, using 10-cm VAS, the relative per cent change is difficult to calculate if there is a change from 0 to a higher value.

The myositis experts decided to use similar response criteria for adult DM/PM and juvenile DM, to facilitate combined clinical trials, such as the RIM trial.[3] Another advantage of the response criteria is that although they are the same for adult DM/PM and juvenile DM, they address the unique differences in the core set measure responsiveness between the 2 disease entities by specifying higher thresholds for juvenile DM than for adult DM/PM, which reflects the fact that more responsiveness is seen in juvenile DM patients in clinical trials.[3 5] Additionally, the juvenile DM response criteria allow for the possibility of using the IMACS or PRINTO core set measures and provide a more definitive threshold for major improvement.[20]

Some limitations of the new response criteria should be noted. First, most of the core set measures, although proven to have good reliability and validity, are subjective and evaluator dependent. However, similar metrics have been used successfully in RA trials that used a physician global measure similar to that used for myositis.

Second, only one major clinical trial was available for validation, and it failed to meet its primary end point and was not truly placebo controlled. Thus, we validated the results using the treating physician's improvement scores in the clinical trial.

Third, the threshold for major improvement in the response criteria is considered preliminary due to an insufficient number of adult DM/PM cases showing major improvement. We believe that future studies using therapeutic agents that have a greater impact on myositis disease activity will lead to better clinical responses, thus allowing investigators to determine a final threshold for major improvement. We plan to validate major improvement in future studies.

Fourth, given that the criteria are focused on improvement and thus fail to differentiate between no change and worsening, these criteria might not be applicable in studies of worsening disease activity (ie, disease flare designs) in myositis. However, in the future, it will be necessary to develop criteria for flare in myositis.

Fifth, the response criteria were developed using a PM diagnosis based on the Bohan and Peter classification criteria, but experts now recognise that PM, according to those criteria, may include different syndromes, such as necrotising myopathy, the antisynthetase syndrome, and others.[42 43] We believe that these response criteria will still be applicable to these newer entities given that the data- and consensus-driven processes described herein were inclusive of those syndromes. In the future, with changes in classification criteria terminology,[44] the response criteria terminology will need to be modified accordingly.

Sixth, because the criteria are complex and might be difficult to apply in research studies, we are developing a web-based tool as well as a downloadable calculator that will allow easy application of the response criteria. The time required to apply these criteria is estimated to be 25 min to complete the core set measures at each visit[6] and 3 min to hand-calculate the total improvement score and degree of response, while with a computer-based system the calculation time is negligible. Moreover, although the criteria may appear to be complicated, the core set measures to be collected by any study or investigators are simple and are essentially the same as those in previous myositis studies and trials.

Finally, patient-reported outcomes as core set measures, with the exception of the HAQ and patient global assessment, were not part of the response criteria, perhaps due to the paucity of sensitive and responsive patient-reported outcomes for DM/PM.[45]

In conclusion, the development of data- and consensus-driven conjoint analysis-based continuous response criteria with quantitative assessment of improvement on a scale of 0–100 and with thresholds for minimal, moderate, and major (preliminary threshold) improvement marks a major advancement in assessing response in myositis clinical trials and studies. These response criteria are sensitive and specific and provide a way to determine clinically meaningful change corresponding to degree of clinical improvement. These response criteria were valid in a clinical trial and had excellent face validity and acceptance among myositis experts from various specialties who care for adult DM/PM patients in different parts of the world. A conjoint analysis-based definition with a continuous improvement score using absolute per cent change in core set measures with thresholds for minimal, moderate, and major improvement was selected as the response criteria to be used for adult clinical trials.

Author affiliations
[1]University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[2]NIEHS, NIH, Bethesda, Maryland, USA
[3]Istituto Giannina Gaslini, Pediatria II - Rheumatologia, PRINTO, Genoa, Italy
[4]Social and Scientific Systems, Inc., Durham, North Carolina, USA
[5]The Hospital for Sick Children, Toronto, Ontario, Canada
[6]Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA
[7]Central Manchester University Hospitals NHS Foundation Trust, University of Manchester, Manchester, UK
[8]University of Liverpool, Liverpool, UK
[9]Karolinska University Hospital, Karolinska Institute, Stockholm, Sweden
[10]Stanford University, Redwood City, California, USA
[11]University College London, London, UK
[12]Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[13]Academic Medical Center, Amsterdam, The Netherlands
[14]Mayo Clinic, Rochester, Minnesota, USA
[15]University of Debrecen, Debrecen, Hungary
[16]Hospital General de Occidente de la Secretaría de Salud and University of Guadalajara, Guadalajara, México
[17]Graduate School of Convergence Science and Technology and Seoul National University Hospital, Seoul, Korea
[18]Charles University, Prague, Czech Republic

Twitter Follow Hector Chinoy @drhectorchinoy

Acknowledgements We thank the following individuals for providing invaluable input and feedback on project development and support: members of the American College of Rheumatology Criteria Committee; Dr Daniel Aletaha (European League Against Rheumatism), Drs Suzette Peng and Sarah Yim (US Food and Drug Administration), Drs Thorsten Vetter and Richard Vesely (European Medicines Agency), Bob Goldberg and Theresa Curry (The Myositis Association), Rhonda McKeever and Patti Lawler (Cure JM Foundation), and Irene Oakley (Myositis UK). We also thank Drs Michael Ward, Steven Pavletic, and Adam Schiffenbauer for their critical review of the manuscript. Paul Hansen, who with Franz Ombler owns and co-invented the 1000Minds software referred to in the article, provided intellectual and logistic support for this project.

Collaborators APPENDIX A: MEMBERS OF THE INTERNATIONAL MYOSITIS ASSESSMENT AND CLINICAL STUDIES GROUP AND THE PAEDIATRIC RHEUMATOLOGY INTERNATIONAL TRIALS ORGANISATION WHO CONTRIBUTED TO DEVELOPING THE RESPONSE CRITERIA:
Steering committee: Lisa G Rider (co-principal investigator), Nicolino Ruperto (co-principal investigator), Rohit Aggarwal (methodology lead), Frederick W Miller, Jiri Vencovsky.
Statistical team: Rohit Aggarwal, Brian Erman, Nastaran Bayat, Angela Pistorio, Adam M Huber, Brian M Feldman, Paul Hansen, Howard Rockette, Peter A Lachenbruch, Nicolino Ruperto, Lisa G Rider.
Adult core set survey group: Anthony A Amato, Hector Chinoy, Lisa Christopher-Stine, Lorinda Chung, Robert G Cooper, Lisa Criscione-Schreiber, Leslie Crofford, Mary E Cronin, Katalin Dankó, David Fiorentino, Ignacio García-De la Torre, Patrick Gordon, Gerald Hengstman, James D Katz, Andrew Mammen, Galina Marder, Neil McHugh, Chester V Oddis, Elena Schiopu, Albert Selva-O'Callaghan,

Yeong Wook Song, Jiri Vencovsky, Gil Wolfe, Robert Wortmann.
Clinical trial or natural history study data set contributions: Anthony A Amato, Hector Chinoy, Lorinda Chung, Robert G Cooper, Katalin Dankó, David Fiorentino, Ignacio García-De la Torre, Mark Gourley, Ingrid Lundberg, Frederick W Miller, Chester V Oddis, Paul Plotz, Lisa G Rider, Yeong Wook Song, Jiri Vencovsky.
Adult patient profile working group: Rohit Aggarwal, Anthony A Amato, Dana Ascherman, Richard Barohn, Olivier Benveniste, Jan De Bleecker, Jeffrey Callen, Christina Charles-Schoeman, Hector Chinoy, Lisa Christopher-Stine, Lorinda Chung, Robert G Cooper, Leslie Crofford, Mary E Cronin, Katalin Dankó, Sonye Danoff, Maryam Dastmalchi, Marianne de Visser, Mazen Dimachkie, Steve DiMartino, Lyubomir Dourmishev, Floranne Ernste, David Fiorentino, Ignacio García-De la Torre, Takahisa Gono, Patrick Gordon, Mark Gourley, David Isenberg, Yasuhiro Katsumata, James D Katz, John Kissel, Richard L Leff, Todd Levine, Ingrid Lundberg, Andrew Mammen, Herman Mann, Galina Marder, Isabelle Marie, Neil McHugh, Joseph Merola, Frederick W Miller, Chester V Oddis, Marzena Olesinska, Nancy Olsen, Nicolo Pipitone, Sindhu Ramchandren, Seward Rutkove, Lesley Ann Saketkoo, Adam Schiffenbauer, Albert Selva-O'Callaghan, Samuel Katsuyuki Shinjo, Rachel Shupak, Yeong Wook Song, Katarzyna Swierkocka, Jiri Vencovsky, Julia Wanschitz, Victoria Werth, Irene Whitt, Robert Wortmann, Steven R Ytterberg.
Conjoint analysis, adult group: Rohit Aggarwal, Anthony A Amato, Hector Chinoy, Lisa Christopher-Stine, Lorinda Chung, Robert G Cooper, Mary E Cronin, Katalin Dankó, Mazen Dimachkie, Steve DiMartino, David Fiorentino, Ignacio García-De la Torre, Patrick Gordon, Ingrid Lundberg, Herman Mann, Frederick W Miller, Chester V Oddis, Albert Selva-O'Callaghan, Jiri Vencovsky, Victoria Werth, Robert Wortmann, Steven R Ytterberg.
Participants in consensus conference, adult working group: Anthony A Amato, Hector Chinoy, Robert G Cooper, Maryam Dastmalchi, Marianne de Visser, David Fiorentino, David Isenberg, James D Katz, Andrew Mammen, Chester V Oddis, Jiri Vencovsky, Steven R Ytterberg.
Participants in consensus conference, paediatric working group: Rolando Cimaz, Rubén Cuttica, Sheila Knupp Feitosa de Oliveira, Brian M Feldman, Adam M Huber, Carol B Lindsley, Clarissa Pilkington, Marilynn Punaro, Angelo Ravelli, Ann Reed, Kelly Rouster-Stevens, Annet van Royen-Kerkhof.

Contributors 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 Aggarwal and Rider 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. All authors: Study conception and design; Acquisition of data; and Analysis and interpretation of data.

Funding Supported in part by the American College of Rheumatology, the European League Against Rheumatism, Cure JM Foundation, Myositis UK, Istituto G. Gaslini and the Paediatric Rheumatology International Trials Organisation (PRINTO), the Myositis Association, and the NIH (National Institute of Environmental Health Sciences (NIEHS), National Center for Advancing Translational Sciences, and National Institute of Arthritis and Musculoskeletal and Skin Diseases). IG-DIT's work was supported in part by CONACYT (Programa Nacional de Posgrados de Calidad). YWS's work was supported by the Korea Health Technology R & D Project through the Korea Health Industry Development Institute funded by the Ministry of Health & Welfare, Republic of Korea (grant HI14C1277). JV's work was supported by the Ministry of Health, Czech Republic (Institute of Rheumatology project for conceptual development of a research organisation, 00023728).

Competing interests None declared.

Provenance and peer review Not commissioned; internally peer reviewed.

REFERENCES

1 Rider LG, Miller FW. Deciphering the clinical presentations, pathogenesis, and treatment of the idiopathic inflammatory myopathies. *JAMA* 2011;305:183–90.
2 Moghadam-Kia S, Aggarwal R, Oddis CV. Treatment of inflammatory myopathy: emerging therapies and therapeutic targets. *Expert Rev Clin Immunol* 2015;11:1265–75.
3 Oddis CV, Reed AM, Aggarwal R, *et al*. Rituximab in the treatment of refractory adult and juvenile dermatomyositis and adult polymyositis: a randomized, placebo-phase trial. *Arthritis Rheum* 2013;65:314–24.
4 Muscle Study Group. A randomized, pilot trial of etanercept in dermatomyositis. *Ann Neurol* 2011;70:427–36.
5 Ruperto N, Pistorio A, Oliveira S, *et al*. Prednisone versus prednisone plus ciclosporin versus prednisone plus methotrexate in new-onset juvenile dermatomyositis: a randomised trial. *Lancet* 2016;387:671–8.
6 Rider LG, Werth VP, Huber AM, *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 (IBMFRS), 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.
7 Miller FW, Rider LG, Chung YL, *et al*. Proposed preliminary core set measures for disease outcome assessment in adult and juvenile idiopathic inflammatory myopathies. *Rheumatology (Oxford)* 2001;40:1262–73.
8 Rider LG, Giannini EH, Harris-Love M, *et al*. Defining clinical improvement in adult and juvenile myositis. *J Rheumatol* 2003;30:603–17.
9 Rider LG, Giannini EH, Brunner HI, *et al*. International consensus on preliminary definitions of improvement in adult and juvenile myositis. *Arthritis Rheum* 2004;50:2281–90.
10 American College of Rheumatology Committee to Reevaluate Improvement Criteria. A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. *Arthritis Rheum* 2007;57:193–202.
11 De Lautour H, Taylor WJ, Adebajo A, *et al*. Development of preliminary remission criteria for gout using Delphi and 1000Minds consensus exercises. *Arthritis Care Res (Hoboken)* 2016;68:667–72.
12 Johnson SR, Naden RP, Fransen J, *et al*. Multicriteria decision analysis methods with 1000Minds for developing systemic sclerosis classification criteria. *J Clin Epidemiol* 2014;67:706–14.
13 Hansen P, Ombler F. A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *J Multi Crit Decis Anal* 2008;15:87–107.
14 Amaya-Amaya M, Gerard K, Ryan M. Discrete choice experiments in a nutshell. In: Ryan M, Gerard K, Amaya-Amaya M, eds. *Using discrete choice experiments to value health and health care*. Dordrecht: Springer, 2008.
15 Ruperto N, Martini A. Networking in paediatrics: the example of the Paediatric Rheumatology International Trials Organisation (PRINTO). *Arch Dis Child* 2011;96:596–601.
16 Taylor WJ, Singh JA, Saag KG, *et al*. Bringing it all together: a novel approach to the development of response criteria for chronic gout clinical trials. *J Rheumatol* 2011;38:1467–70.
17 Taylor WJ, Brown M, Aati O, *et al*. Do patient preferences for core outcome domains for chronic gout studies support the validity of composite response criteria? *Arthritis Care Res (Hoboken)* 2013;65:1259–64.
18 Utz KS, Hoog J, Wentrup A, *et al*. Patient preferences for disease-modifying drugs in multiple sclerosis therapy: a choice-based conjoint analysis. *Ther Adv Neurol Disord* 2014;7:263–75.
19 de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ* 2012;21:145–72.
20 Rider LG, Aggarwal R, Pistorio A, *et al*. 2016 American College of Rheumatology/European League Against Rheumatism criteria for minimal, moderate and major clinical response for juvenile dermatomyositis: an International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation collaborative initiative. *Arthritis Rheumatol*. http://onlinelibrary.wiley.com/doi/10.1002/art.40060/abstract
21 Fries JF, Spitz P, Kraines RG, *et al*. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
22 Rider LG. Outcome assessment in the adult and juvenile idiopathic inflammatory myopathies. *Rheum Dis Clin North Am* 2002;28:935–77.
23 Aggarwal R, Rider LG, Ruperto N, *et al*. A consensus hybrid definition using a conjoint analysis is the proposed response criteria for minimal and moderate improvement for adult polymyositis and dermatomyositis clinical trials [abstract]. *Arthritis Rheumatol* 2014;66:S404.
24 Rider LG, Ruperto N, Pistorio A, *et al*. 2016 development of adult dermatomyositis and polymyositis and juvenile dermatomyositis response criteria: methodological aspects: an American College of Rheumatology/European League Against Rheumatism/International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation collaborative initiative. *Rheumatology (Oxford)*. (in press).
25 Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med* 1975;292:344–7.
26 Rider LG, Lee J, Jansen A, *et al*. Defining clinically relevant changes in core set activity measures for adult and juvenile idiopathic inflammatory myopathies (IIM). *Arthritis Rheum* 2011;63:S89.
27 Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283–98.
28 Ruperto N, Meiorin S, Iusan SM, *et al*. Consensus procedures and their role in pediatric rheumatology. *Curr Rheumatol Rep* 2008;10:142–6.
29 Delbecq A, van de Ven A, Gustafson D. *Group techniques for program planning: a guide to nominal group and Delphi processes*. Glenview, IL: Scott, Foresman and Company, 1975.
30 Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. *Semin Arthritis Rheum* 2011;41:95–105.
31 Ruperto N, Ozen S, Pistorio A, *et al*. EULAR/PRINTO/PRES criteria for Henoch-Schonlein purpura, childhood polyarteritis nodosa, childhood Wegener granulomatosis

and childhood Takayasu arteritis: Ankara 2008. Part I: overall methodology and clinical characterisation. *Ann Rheum Dis* 2010;69:790–7.
32 Piram M, Koné-Paut I, Lachmann HJ, *et al*. Validation of the auto-inflammatory diseases activity index (AIDAI) for hereditary recurrent fever syndromes. *Ann Rheum Dis* 2014;73:2168–73.
33 Ware JE Jr, Snow KK, Kosinski M, *et al*. *SF-36 health survey: manual and interpretation guide*. Boston: The Health Institute, New England Medical Center, 1993.
34 Ruperto N, Ravelli A, Pistorio A, *et al*. The provisional Paediatric Rheumatology International Trials Organisation/American College of Rheumatology/European League Against Rheumatism Disease activity core set for the evaluation of response to therapy in juvenile dermatomyositis: a prospective validation study. *Arthritis Rheum* 2008;59:4–13.
35 Ruperto N, Ravelli A, Pistorio A, *et al*. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries: review of the general methodology. *Clin Exp Rheumatol* 2001;19(Suppl 23):S1–9.
36 Apaz MT, Saad-Magalhães C, Pistorio A, *et al*. Health-related quality of life of patients with juvenile dermatomyositis: results from the Paediatric Rheumatology International Trials Organisation Multinational Quality of Life cohort study. *Arthritis Rheum* 2009;61:509–17.
37 Van den Hoogen F, Khanna D, Fransen J, *et al*. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2013;65:2737–47.
38 Neogi T, Aletaha D, Silman AJ, *et al*. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: phase 2 methodological report. *Arthritis Rheum* 2010;62:2582–91.
39 Neogi T, Jansen TL, Dalbeth N, *et al*. 2015 gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheumatol* 2015;67:2557–68.
40 Felson DT, Anderson JJ, Boers M, *et al*. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38:727–35.
41 Felson DT, Smolen JS, Wells G, *et al*. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum* 2011;63:573–86.
42 Aggarwal R, Cassidy E, Fertig N, *et al*. Patients with non-Jo-1 anti-tRNA-synthetase autoantibodies have worse survival than Jo-1 positive patients. *Ann Rheum Dis* 2014;73:227–32.
43 Hengstman GJ, ter Laak HJ, Vree Egberts WT, *et al*. Anti-signal recognition particle autoantibodies: marker of a necrotising myopathy. *Ann Rheum Dis* 2006;65:1635–8.
44 Tjarnlund A, Bottai M, Rider LG, *et al*. Progress report on development of classification criteria for adult and juvenile idiopathic inflammatory myopathies [abstract]. *Arthritis Rheum* 2012;64:S323–4.
45 Alexanderson H, del Grande M, Bingham CO III, *et al*. Patient-reported outcomes and adult patients' disease experience in the idiopathic inflammatory myopathies. Report from the OMERACT 11 Myositis Special Interest Group. *J Rheumatol* 2014;41:581–92.
46 Volochayev R, Csako G, Wesley R, *et al*. Laboratory test abnormalities are common in polymyositis and dermatomyositis and differ among clinical and demographic groups. *Open Rheumatol J* 2012;6:54–63.

EULAR/EFORT recommendations for management of patients older than 50 years with a fragility fracture and prevention of subsequent fractures

W F Lems,[1] K E Dreinhöfer,[2] H Bischoff-Ferrari,[3] M Blauth,[4] E Czerwinski,[5] JAP da Silva,[6] A Herrera,[7] P Hoffmeyer,[8] T Kvien,[9] G Maalouf,[10] D Marsh,[11] J Puget*,[12] W Puhl,[13] G Poor,[14] L Rasch,[1] C Roux,[15] S Schüler,[2] B Seriolo,[16] U Tarantino,[17] T van Geel,[18] A Woolf,[19] C Wyers,[20,21] P Geusens[22,23]

Handling editor Hans WJ Bijlsma

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210289).

For numbered affiliations see end of article.

Correspondence to
Professor Willem Lems, Department of Rheumatology, Amsterdam Rheumatology and Immunology Center, VU University Medical Center, Postbox 7057, 1007 MB Amsterdam, The Netherlands wf.lems@vumc.nl

WFL and KED are co-first authors.

*Jean Puget has since passed away.

Received 31 July 2016
Revised 13 November 2016
Accepted 2 December 2016
Published Online First 22 December 2016

ABSTRACT

The European League Against Rheumatism (EULAR) and the European Federation of National Associations of Orthopaedics and Traumatology (EFORT) have recognised the importance of optimal acute care for the patients aged 50 years and over with a recent fragility fracture and the prevention of subsequent fractures in high-risk patients, which can be facilitated by close collaboration between orthopaedic surgeons and rheumatologists or other metabolic bone experts. Therefore, the aim was to establish for the first time collaborative recommendations for these patients. According to the EULAR standard operating procedures for the elaboration and implementation of evidence-based recommendations, 7 rheumatologists, a geriatrician and 10 orthopaedic surgeons met twice under the leadership of 2 convenors, a senior advisor, a clinical epidemiologist and 3 research fellows. After defining the content and procedures of the task force, 10 research questions were formulated, a comprehensive and systematic literature search was performed and the results were presented to the entire committee. 10 recommendations were formulated based on evidence from the literature and after discussion and consensus building in the group. The recommendations included appropriate medical and surgical perioperative care, which requires, especially in the elderly, a multidisciplinary approach including orthogeriatric care. A coordinator should setup a process for the systematic investigations for future fracture risk in all elderly patients with a recent fracture. High-risk patients should have appropriate non-pharmacological and pharmacological treatment to decrease the risk of subsequent fracture.

INTRODUCTION

Osteoporosis is the most common cause of fragility fractures. These fractures—most frequently occurring at the hip, vertebra, proximal humerus and distal radius—are associated with an increased morbidity and mortality and have a large medical and economic impact on healthcare systems.[1]

Fragility fractures in women and men older than 50 years are among the most frequent musculoskeletal manifestations for which patients consult healthcare providers from more than one medical specialty. Immediately following a fracture, the patient needs acute fracture care, supplied by an orthopaedic or trauma surgeon, and perioperative medical care for the, often fragile, patient. This is followed by the implementation of fracture prevention modalities in patients at risk for a subsequent fracture. This is usually executed under the supervision of general practitioners, rheumatologists or other metabolic bone disease experts. Obviously, a close collaboration between these specialties is necessary at a local level.

Both the European League Against Rheumatism (EULAR) and the European Federation of National Associations of Orthopaedics and Traumatology (EFORT) have recognised the importance of optimal multidisciplinary care for patients with a recent fracture, followed by prevention of subsequent fractures in high-risk patients, and have therefore collaboratively initiated this recommendation.

METHODS

This is the second combined task force for EULAR/EFORT: in line with the first combined recommendations on the swollen knee[2] the EULAR standardised operating procedures for the elaboration and implementation of evidence-based recommendations[3] were initially followed and later updated, when possible, to the 2014 update of the recommendations.[4]

The executive committee comprised the convenors (KD invited by EFORT, WL invited by EULAR), a senior advisor (PG), a clinical epidemiologist (CW) and three research fellows (SS, LR, TvG). Subsequently, the executive committee invited 7 rheumatologists from 7 countries and 10 orthopaedic surgeons from 10 countries selected on the basis of their field of interest and knowledge, while allowing for a broad coverage in the field ensuring in their selections an appropriate geographic distribution of experts across Europe.

During the first group meeting, we started with a general discussion about the management of the patient with an acute fracture and subsequent fracture prevention, and asked all committee members to bring up 10 propositions for research questions. Consensus on the research questions was reached following the Delphi technique. We started with a list of all proposals; overlapping propositions were merged. The list was sent to the experts and they were asked to select the 10 most important propositions from the list. Propositions were accepted automatically if selected by over half of participants in any round, whereas propositions receiving three votes or less were removed. The other propositions

To cite: Lems WF, Dreinhöfer KE, Bischoff-Ferrari H, *et al*. *Ann Rheum Dis* 2017;**76**:802–810.

entered the subsequent Delphi round. The procedure was performed 3 times until we had 15 propositions, which were merged by KD, PG and WFL into the 10 final research questions, as a base for formulating recommendations. In total, three Delphi rounds, facilitated by the convenors, were performed by email.

A systematic literature research (SLR), based on these 10 research questions, was undertaken by the research fellows supported by their mentors in three groups: LR/WL, SS/KD and TvG/PG and the epidemiologist (CW), in Medline and the Cochrane Database of Systematic Reviews (2000–2014). Study designs of interest were systematic review/meta-analysis, randomised controlled trials (RCT)/controlled trials (CT) and observational studies. For every recommendation, all results obtained by the research fellows were discussed with the convenors. First all titles and afterwards all abstracts were scanned for relevance: studies that were clearly out of the scope of the SLR were rejected by the research fellows. Studies that were clearly within or doubtful within the SLR were discussed with the convenors. For every recommendation, all results obtained by the research fellows were discussed with the convenors.

Data from the literature reviews were categorised and presented at the second taskforce meeting according to study design, using a hierarchy of evidence in descending order according to quality. The results were presented and broadly discussed at the second meeting. In addition, these results were the starting point for discussions within the committee, finally leading to consensus about 10 recommendations (table 1). Recommendations were developed and circulated to all members three times in total, to achieve consensus on the final formulation of recommendations.

The level of evidence for each recommendation was rated according to the EULAR standard operating procedures (4) and by the Oxford Levels of Evidence, which define the level of evidence based on the type of research (see online supplementarytable S1).The strength of each recommendation is defined by a combination of the information from the SLR (categories of evidence) and expert opinion (see online supplementary table S2).

Finally, every member of the task force had to indicate the level of agreement with each recommendation. This was scored on a numerical rating scale ranging from 0 (completely disagree) to 10 (completely agree). The average, the median and the range have been calculated (see online supplementary table S1).

Finally, after receiving feedback from the EFORT and EULAR boards, the recommendations were again adapted and circulated to the expert group for feedback and agreement. Key publications which appear after the literature search in 2015 were added to the manuscript.

In the 2014 update of the EULAR recommendations the subtle, but important, differences between recommendations and points to consider were discussed.[4] Since the majority, although not all, of the 10 recommendations had evidence-based answers, we mention them as a set of recommendations, and not as 'points to consider', as proposed recently by the EULAR.[4]

RESULTS

The combined search from the systemic literature review for Q1–10 identified a number of articles for each research question, as shown in online supplementary table S3. Articles that were relevant to >1 research question were included in the review more than once. The 10 recommendations, level of evidence, strength of recommendation and the level of agreement are presented in table 1.

RECOMMENDATIONS

Recommendation 1: preoperative and perioperative management

Fragility fractures should be managed in the context of a multidisciplinary clinical system, guaranteeing adequate preoperative assessment and preparation of patients including adequate pain relief, appropriate fluid management and surgery within 48 hours of injury.

Patients with fragility fractures often have pre-existing chronic diseases, which will have an influence on their general management, short-term and long-term survival rate and their functional recovery. Minimising delirium and avoiding complications is critical for achieving good outcomes. Rapid optimisation of fitness for surgery and early surgery seem to improve morbidity and mortality.

Appropriate pain management should be provided to every patient as soon as possible and before starting diagnostic investigations.[5] A meta-analysis has demonstrated that the use of nerve blocks reduces acute pain in patients suffering from a hip fracture.[6]

The systematic multidisciplinary and comprehensive admission assessment of the patient's medical conditions should include investigations for the most common modifiable variables: malnutrition, electrolyte or volume disturbances, anaemia, cardiac or pulmonary diseases, dementia and delirium and glycaemic control.[7–10] Preoperative investigations should include chest X-ray, ECG, full blood count, clotting studies, blood group, renal function, in addition assessment of cognitive baseline function. This should allow identification and treatment of exacerbations of chronic medical conditions or acute medical illness when appropriate.[11]

Safe and timely transfer from the emergency room to an orthogeriatric ward and definitive treatment including early surgery within 24–48 hours after admission significantly reduces short-term and mid-term mortality rates[12] and reduces minor and major medical complications due to immobility and its accompanying effects (eg, decubitus ulcer, pneumonia, increased length of hospital stay).[9 13 14] Delay to the operation theatre to enable optimisation of acute medical problems has to be weighed up against the effects of prolonging pain and immobility.

Recommendation 2: orthogeriatric care

To improve functional outcome, and to reduce length of hospital stay and mortality, orthogeriatric comanagement should be provided, especially in elderly patients with hip fracture.

Elderly fracture patients admitted to the hospital will benefit from multidisciplinary comanagement, including a comprehensive geriatric assessment[15] of medical, functional and psychological capabilities and adequate preparation before surgery.[16 17] In patients with hip fracture, the joint care model between geriatrician and orthopaedic surgeon on a dedicated orthogeriatric ward has been shown to have the shortest time to surgery, the shortest length of inpatient stay and the lowest inpatient and 1-year mortality rate.[18–20]

Patients with fragility fractures are at risk for multiple postoperative complications: some are patient related, while others are related to the surgical treatment. In the elderly multimorbid patient, complications are frequent and may increase the length of stay and perioperative mortality.[21] Complications are related to increased mortality and morbidity, and therefore should

Table 1 Recommendations for patients with fragility fractures in patients aged 50 years and older

Recommendation		Level of evidence	Strength of recommendation	Level of agreement Average Median Range
1	Fragility fractures should be managed in the context of a multidisciplinary clinical system, guaranteeing adequate preoperative assessment and preparation of patients, including adequate pain relief, appropriate fluid management and surgery within 48 hours of injury	IIA	B	9.8 10 8–10
2	To improve functional outcome, and to reduce length of hospital stay and mortality, orthogeriatric comanagement should be provided, especially in elderly patients with hip fracture	IA	A	9.2 10 0–10
3	Appropriate treatment of the fractures in these, often elderly and multimorbid, patients with frail bones requires a balanced approach with regard to operative vs non-operative treatment and careful selection of fixation devices and techniques	III	C	9.3 10 7–10
4	Each patient aged 50 years and over with a recent fracture should be evaluated systematically for the risk of subsequent fractures	IA	A	9.5 10 5–10
5	Evaluation of the risk of subsequent fractures includes a review of clinical risk factors, DXA of the spine and hip, imaging of the spine for vertebral fractures and evaluation of falls risk and the identification of secondary osteoporosis, which together predict subsequent fracture risk	III	C	9.3 10 6–10
6	Implementation requires a local responsible lead, that is, a person/group that coordinates secondary fracture prevention based on guidelines, liaising between surgeons, rheumatologists/endocrinologists, geriatricians in case of elderly with a hip or other major fracture, and general practitioners	IV	D	9.1 10 6–10
7	An appropriate rehabilitation programmes should consist of both early postfracture introduction of physical training and muscle strengthening and the long-term continuation of balance training and multidimensional fall prevention	IIA	B	9.5 10 5–10
8	Patients should be educated about the burden of the disease, risk factors for fractures, follow-up and duration of therapy	IV	D	9.2 10 5–10
9	Non-pharmacological treatment is important in the prevention of fractures in high-risk patients; it includes at least an adequate intake of calcium and vitamin D, stopping smoking and limitation of alcohol intake	IV	D	9.3 10 6–10
10	Pharmacological treatment should preferably use drugs that have been demonstrated to reduce the risk of vertebral, non-vertebral and hip fractures, and should be regularly monitored for tolerance and adherence	IB	A	9.9 10 9–10

DXA, dual energy xray absorptiometr.

preferably be prevented, if possible: delirium,[22] deep venous thrombosis,[23–25] pressure sores[26] and malnutrition.[27 28]

Postoperative care should include appropriate pain management and antibiotic prophylaxis, correction of postoperative anaemia, routine systems examinations, regular assessment of cognitive function, assessment for pressure sores, nutritional status and renal function, assessment and regulation of bowel and bladder function, wound assessment and care and early mobilisation.[20]

Recommendation 3: treatment of the fracture

Appropriate treatment of the fractures in these often elderly and multimorbid patients with frail bones requires a balanced approach with regard to operative versus non-operative treatment and careful selection of fixation devices and techniques.

Recommendations for surgical treatment are of course dependent on the type of fracture and on the individual patient.[29]

Distal radius fracture

Distal radius fractures after a fall from standing height can be treated by cast immobilisation or by operative methods including locking plates, Kirschner wires or external fixation. Recent RCTs have not identified clear recommendations for the optimal treatment in the elderly population.[30–32] In a systematic review cast immobilisation had the worst radiographic outcome but the least complications and a comparable functional outcome with surgical treatment options.[31] Radiographic alignment after closed reduction and the functional demand of the patient should guide the decision for further operative stabilisation.[29]

Vertebral fractures

Only one out of three vertebral fragility fractures are symptomatic and about 10% of patients will require hospitalisation because of pain. Most symptomatic fractures are treated with analgesics, activity modification and bracing,[33 34] and so far there are inconclusive results on surgical versus non-surgical interventions.[35–38]

Hip fractures

Hip fractures are common, have often devastating effects on the patients and usually require surgical intervention. Treatment options are depended on fracture location and classification, age, functional status of the patient and pre-existing osteoarthritis.

Femoral neck fractures

Stable non-displaced fractures can be addressed with cannulated screw fixation in a percutaneous manner.[39] Displaced femoral neck fractures in healthy, active and independent older individuals without cognitive dysfunction are best treated by total hip

arthroplasty allowing immediate full weight-bearing.[40 41] In frail patients, hemiarthroplasty might be preferred, since operative time is shorter and the subsequent dislocation risk is lower while the functional outcome is acceptable.[42] Total hip arthroplasty may offer improved function and long-term results,[43] but patient factors and surgeon experience need to be considered in order to justify the risk of a more complex and costly procedure.[13]

Trochanteric fractures

For stable intertrochanteric fractures a sliding hip screw is favoured, unstable intertrochanteric fractures are treated with an antegrade cephalomedullary nail. Strong evidence supports that cephalomedullary devices should also to be used in subtrochanteric or reverse oblique fractures.[44]

Humerus fractures

Most proximal humeral fractures can be treated non-operatively with good functional outcomes. Treatment of displaced three-part and four-part fractures remains controversial: open reduction and locking plate osteosynthesis is associated with considerable complication, the outcome of hemiarthroplasty is closely related to tuberosity healing. Reverse shoulder arthroplasty may provide satisfactory shoulder function in geriatric patients with pre-existing rotator cuff dysfunction or after the failure of first-line treatment.[45–47]

Recommendation 4: organisation of postfracture care

Each patient aged 50 years and over with a recent fracture should be evaluated systematically for the risk of subsequent fractures.

Since the treatment gap is high, many programmes have been developed to address secondary fracture prevention.[48] The simplest form of intervention is to provide only specific patient education; a more elaborate scheme is alerting the primary care physician (PCP) by means of a discharge letter containing medical information on the fracture of the patient. However, a systematic review and meta-analysis has shown that the Fracture Liaison Service (FLS) is the most effective organisational structure for risk evaluation and treatment initiation.[49]

The central element of an FLS model is a dedicated coordinator who takes care of all aspects of the process (identification, investigation and intervention with therapy).[50] The coordinator is often a well-educated nurse, who works under supervision of an orthopaedic surgeon, an endocrinologist or a rheumatologist. The coordinator is responsible for the identification of all elderly patients with a recent fracture in the hospital, to organise the diagnostic investigations and to start interventions and providing adequate medical information to patients and PCPs.[48]

RCTs[51–53] proved that a nominated coordinator significantly improves the implementation of osteoporosis treatment after a fragility fracture, for example, in a cluster RCT within 6 months after the fracture 45% of patients received appropriate management, while in the control group only 26%.[51]

Recommendation 5: evaluation of subsequent fracture risk

Evaluation of the risk of subsequent fractures includes a review of clinical risk factors, DXA of spine and hip, imaging of the spine for vertebral fractures, evaluation of falls risk and the identification of secondary osteoporosis, which together predict subsequent fracture risk.

Secondary fracture risk is high immediately after the fracture, and gradually decreases over time. Our expert opinion is that in most FLS, patients with fractures 3–6 months before are receiving diagnostic investigations, but investigations at a later stage might also be worthwhile.

Box 1 Tools for evaluation of subsequent fracture risk after an initial fracture

- Clinical risk factors for further fractures:
  - fracture location and severity
  - suboptimal preoperative, operative and postoperative phase with complications and suboptimal rehabilitation
  - high age, low body mass index, personal and family history of fracture, diseases, medications and lifestyle (smoking, alcohol, lack of exercise)
  - fall risk
- DXA of lumbar spine and hips
- Imaging of the spine, by vertebral fracture assessment or by conventional radiographs
- Screening for underlying secondary osteoporosis or other metabolic bone diseases

NB: Clinical risk factors can be integrated in FRAX, Garvan or Q-Fracture algorithms to estimate future fracture risk.

Fracture risk evaluation is recommended to inform therapeutic decisions regarding the prevention of subsequent fractures prevention in high-risk patients[54 55] (box 1).

Apart from the recent fracture location and severity, perioperative complications and suboptimal rehabilitation, clinical risk factors such as advanced age, female gender, low body mass index, lifestyle, personal and family history of fracture, and falls risk all play an important role in subsequent fracture risk.[6 56 57] These are included in fracture risk assessment tools such as FRAX,[58] Garvan[59] and Q-Fracture.[60] In some guidelines, these tools are considered sufficient to make treatment decisions when the risk is identified as being high (based on post hoc analyses), but most guidelines and reimbursement criteria include the results of bone mineral density (BMD) and/or a prevalent hip or vertebral fracture for treatment decisions.[54 55 61 62]

DXA of the lumbar spine and hip is the standard method for measuring BMD, and independently contributes to the assessment of fracture risk.[63] Imaging of the spine by radiography or with vertebral fracture assessment (VFA) (a measurement based on additional software on a DXA device which involves lower irradiation than plain radiographs or CT) allows the detection of subclinical vertebral fractures, which are frequent (20%) in patients with a recent non-vertebral fracture.[64] The presence, number and severity of vertebral fractures are related to fracture risk and contribute to therapeutic decisions, independent of BMD and other risks.[65]

Fall risk evaluation starts with history of falls during last year, followed by specific tests when indicated. A limited standard laboratory examination including erythrocyte sedimentation rate, serum calcium, albumin, creatinine and thyroid-stimulating hormone and other tests (such as vitamin D, protein electrophoresis, testosterone in men, etc) when clinically indicated, allows diagnosis of frequently present subclinical disease (in 30%), which increases the risk of fractures.[66]

Recommendation 6: implementation of guidelines

Implementation requires a local responsible lead, that is, a person/group that coordinates secondary fracture prevention based on guidelines liaising between surgeons, rheumatologists/

endocrinologists, geriatricians in case of elderly with a hip or other major fracture and general practitioners.

Implementation of clinical guidelines in routine daily practice is often difficult. Effective implementation should focus on three basic issues: (a) the level of evidence (eg, RCTs), (b) barriers and facilitators and (c) effectiveness of dissemination and implementation strategies.[67]

Several guidelines or recommendations are available for patients with a recent fragility fracture, such as those from American Association of Orthopedic Surgeons (AAOS),[68] British Orthopaedic Association (BOA),[69] American Society of Bone and Mineral Research (ASBMR)[54] and International Osteoporosis Foundation (IOF);[55] however, our recommendations are unique since they are the first that combined recommendations for acute fracture care and for subsequent fracture prevention.

The National Hip Fracture Database initiative was conceived as a clinician-led collaboration between the BOA and the British Geriatrics Society, in which six clinical standards for hip fracture care were agreed.[69] This clinician-led audit initiative has led to substantial improvements in care and survival of older people with hip fracture in England.[70] The implementation of an evidence-based algorithm for hip fracture surgery in Denmark facilitated a low reoperation rate.[71] In the acute fracture care phase, orthogeriatric comanagement are recommended for the frail, elderly patient with multiple comorbidities and polypharmacy[17 18 72] and has been shown to bring about a decreased length of stay[73] and improved mobility.[17]

Implementation of guidelines should adapt to local needs and restrictions and should be based on collaboration between orthopaedic surgeons, rheumatologists/endocrinologists, geriatricians (in case of elderly with a hip or other major fracture) and general practitioners.[18 48 54 55]

Recommendation 7: rehabilitation

An appropriate rehabilitation programme should consist of both the early postfracture introduction of physical training and muscle strengthening and the long-term continuation of balance training and multidimensional fall prevention.

The most important aim for all patients sustaining a fragility fracture is to regain the level of mobility and independence they enjoyed before the fracture occurred. Early identification of individual goals and needs are essential for each patient, before the rehabilitation plan can be developed. Especially in the elderly, a multidisciplinary and multifactorial comprehensive rehabilitation programme is recommended.[74–77]

Early mobilisation following surgery, preferably starting on the first postoperative day, is critical for a patient's functional independence and prevention of postoperative complications.[76]

In patients with hip fracture, this comprises immediate weight bearing,[78] early ambulation[79] as tolerated by the patient and transfer training in and out of bed. Based on the initial condition of the patient, appropriate physical therapy includes upper-extremity and lower-extremity strength exercises, gait training (eg, on a treadmill),[80] balance and functional training (eg, ambulation and stair climbing) as well as aerobic[81] and stretching exercises for tight soft tissues and joints.

For patients with vertebral fractures, a recent Cochrane Review[82] found inconclusive results for the effect of exercise or active physical therapy interventions in these patients and no definitive conclusion could be drawn. Only moderate evidence seems to exist with regard to improvement of walking speed, back extensor strength, trunk muscle endurance, quality of life and pain.

After casting or surgery for distal radius fracture, early finger motion is essential to prevent oedema and stiffness. When immobilisation is discontinued, aggressive finger and hand motion is necessary to facilitate the best possible outcomes.

Following surgical treatment of a fracture of the shoulder, range-of-motion exercises including shoulder, elbow, wrist and hand motion should begin within the first postoperative days. A sling is usually worn for comfort only and may be discarded as early as the patient's pain allows. Above chest level activities should be restricted in the case of both operative and non-operative management until fracture healing is evident. Overly aggressive physical therapy and exercises may increase the risk of fixation failure in the postoperative period.

Exercise programmes and fall prevention programmes are hallmarks of ideal non-pharmacological treatment for the prevention of fractures. Positive effects on BMD and muscle strength are described in patients who exercise rigorously, as well as a reduction in the frequency of falls, but the evidence for fracture prevention is limited.[83]

Recommendation 8: education

Patients should be educated about the burden of the disease, risk factors for fractures, follow-up and duration of therapy.

Perception of fracture risk and the use of BMD testing are higher in patients with a recent fracture when compared with patients without a fracture history.[84]

In RCTs, a systematic review and meta-analyses, written materials with and without video supplements, behavioural frameworks sent out in three mailings for patients, and in patient education to the provider did not affect diagnosis of underlying osteoporosis and subsequent treatment.[48 49 85–87] In a meta-analysis, BMD testing and treatment initiation were lowest in patients who had only education.[87] In a randomised study, a more personalised approach with a phone call plus follow-up letter to patients did not significantly increase osteoporosis follow-up care compared with simply sending out a letter.[88]

Patient education is recommended as an overarching principle and is incorporated in the guidelines as part of fracture prevention programmes.[89]

Recommendation 9: non-pharmacological treatment

Non-pharmacological treatment is important in the prevention of fractures in high-risk patients; it includes at least an adequate intake of calcium and vitamin D, stopping smoking and limitation of alcohol intake.

A non-healthy lifestyle may have negative effects on BMD, bone quality and the risk of falling[83] and should be corrected (stop smoking, limit alcohol intake).

Data on the effects of non-pharmacological treatment on fracture incidence are limited. Calcium and vitamin D were part of the medical treatment in all RCTs, and adequate total calcium intake (diet and when necessary supplementation) of 1000–1200 mg/day together with vitamin D 800 IU/day is advocated when using anti-osteoporosis drugs.

Calcium alone has no demonstrated effect on fracture reduction, and is associated with gastrointestinal side effects, while there is uncertainty whether high calcium intake is associated with cardiovascular events.[90]

Vitamin D deficiency is endemic worldwide, as it is in patients with a recent fracture.[91] Vitamin D supplementation (800 IU/day), with adequate calcium intake, is associated with a 15%–20% reduction in non-vertebral fractures, and also with a

20% reduction in falls.[92–95] High pulse dosages of vitamin D seem to be associated with increased fall risk and fracture risk.[96 97]

Recommendation 10: pharmacological treatment

Pharmacological treatment should preferably use drugs that have been demonstrated to reduce the risk of vertebral, non-vertebral and hip fractures, and should be regularly monitored for tolerance and adherence.

Only one study evaluated the effect of drugs following a recent fracture, namely zoledronic acid, after a recent hip fracture.[98]

Other RCTs have been performed in patients at high risk for subsequent fractures based on the presence of one or more vertebral fractures, and/or a low T-score. Alendronate, risedronate, zoledronic acid (all bisphosphonates) and denosumab (a monoclonal antibody against RANKL) demonstrated a reduction in vertebral fractures, non-vertebral fractures and hip fractures in the primary analyses.[99–102] A reduction in vertebral fractures was demonstrated with raloxifene and ibandronate, and of vertebral and non-vertebral fractures with strontium ranelate and teriparatide.

Alendronate[99] and risedronate[102] are first-choice agents, because these drugs are usually well tolerated, have a low cost (generic forms are available) and physicians may have a lot of experience with oral bisphosphonates. For patients with oral intolerance, dementia, malabsorption and non-compliance zoledronic acid (intravenous)[100] or denosumab (subcutaneous)[101] are alternatives. For patients with very severe osteoporosis, the use of anabolic agents such as teriparatide is an option.[103]

Based on the length of these RCTs, these drugs are usually prescribed for 3–5 years, and longer in patients who remain at high risk. Since long-term adherence to drug treatment is poor, a systematic follow-up is advocated, as part of a five-step plan including identifying patients with a recent fracture: inviting them for fracture risk evaluation; differential diagnosis; therapy and follow-up.[104] Risk communication and shared decision making in the care of patients with osteoporosis may have a positive influence on adherence.[105 106] Adherence to therapy is substantially higher in the FLS (up to 90%), probably because these patients are more motivated because of their recent fracture, and their positive response to an invitation from the FLS.[107]

DISCUSSION

In addition to these recommendations, the group formulated overarching principles that are relevant for optimal care of patients over 50 years of age with a recent fragility fracture.

Overarching principles

First, although both in the acute care phase after the fracture and in the subsequent prevention of secondary fractures, many different medical specialties can be involved, the critical point is not who is taking care of the patient, but that all patients receive optimal care. Obviously, a structured collaboration between healthcare workers is a prerequisite, reflected in several of our recommendations.

Second, optimal acute fracture care is dependent on the type of fracture and the age, presence or absence of comorbidity and the needs of the patient.

Third, especially in the frail elderly person with a major fracture, an orthogeriatric and multidisciplinary approach is warranted.

Fourth, optimal care in the preoperative, operative and postoperative phases has an important effect on clinical outcome. As a consequence, it is very likely that limited mobility and a poor quality of life in the postoperative phase may be associated with an elevated risk of future fractures.

Fifth, for prevention of subsequent fractures, it is important that in all patients fracture risk should be investigated systematically.

Sixth, for subsequent prevention of fractures in high-risk patients, effective and safe drugs should be prescribed, and non-pharmacological treatment options and patient education also need to be considered.

These recommendations and overarching principles can be used as a template for discussions with the local stakeholders (including specialists, general practitioners, fracture nurses, local coordinators, patients and health authorities). Finally, we have included suggestions for further research (box 2).

Box 2 Research agenda

- Factors and interventions that improve the clinical condition of patients with a recent fracture before surgery
- Effects of orthogeriatric assessment on mortality and morbidity in elderly patients with major fractures
- Prevention and treatment of delirium
- Evaluation of the best postfracture rehabilitation strategies for fragility fractures: intensity, duration and content
- Effects of a complex biopsychosocial intervention on early and long-term rehabilitation effects
- Role of muscle loss, sarcopenia and nutrition on recovery following hip fracture, and the role of physical and pharmacological approaches in managing these deficits
- Initiatives for multidisciplinary collaboration for secondary fracture prevention
- What is the long-term effect of fracture liaison service (FLS) and its implementation on adherence to therapy and reduction of fractures, morbidity and mortality
- 'Real-world' cost-effectiveness of orthogeriatric care and for FLS
- Subsequent fracture prevention of individuals who are not able to visit the FLS, for example, patients with hip fracture
- Optimal timing of start and duration of antiosteoporotic drugs
- Benefits of combining exercise, nutrition, pharmacological and other intervention strategies
- Optimise strategies for early fall prevention in patients with fragility fractures
- Effects of drugs (antiresorptive and osteoanabolic drugs, biologics, non-steroidal anti-inflammatory drugs) on fracture healing (delayed or non-union) and on atypical femoral fractures
- Implementation of recommendations.

Limitations

First, the 10 recommendations do not cover all aspects of fragility fracture patient management. Nevertheless, they deal with the main principles of fracture care and secondary fracture prevention, based on the 10 clinical research questions identified by an expert committee. Second, there is a large degree of heterogeneity in patients with a recent fracture, for example, an elderly woman aged 85 years with a hip fracture versus a woman aged 55 years with a wrist fracture. It is understandable that some elderly patients with immobility and comorbidities, as

often seen in patients with a hip or pelvic insufficiency fracture, do not respond to invitations for FLS. For these patients, anti-osteoporotic treatment can be started even without a DXA scan. Third, there is significant heterogeneity of healthcare systems between countries. A fourth limitation is that the scoring of agreement on the level of evidence is best applicable on interventions, but is more difficult to apply to diagnostic procedures. Fifth, we (unfortunately) did not have included a non-medical health professional in the task force. This project started before 2014, and at that time it was not obligatory, and less customary than it is nowadays. Nevertheless, we have described extensively the role that the fracture nurse, as a health professional, could play centrally in the FLS.

CONCLUSION

In conclusion, we provide recommendations for each step of fracture care, which can be integrated into a multidisciplinary approach. This combined EULAR/EFORT task force was characterised by intensive discussions between orthopaedic surgeons and rheumatologists, which strongly increased insight into the thoughts and behaviours of each specialty. We hope that the manuscript will stimulate work between these specialties with fracture patients, both in daily practice and in research projects.

Author affiliations
[1]Department of Rheumatology, Amsterdam Rheumatology and Immunology Center, VU University Medical Center, Amsterdam, The Netherlands
[2]Department of Orthopedics and Traumatology, Center for Musculoskeletal Surgery (CMSC), Charité Universitätsmedizin Berlin Medical Park Berlin Humboldtmühle, Berlin, Germany
[3]Departemnt of Geriatrics and Aging Research, University Hospital and University of Zurich, Zurich, Switzerland
[4]Department for Trauma Surgery, Medical University Innsbruck, Innsbruck, Austria
[5]Department of Bone and Joint Diseases, Jagiellonian University, Faculty of Health and Sciences, Krakow Medical Centre, Krakow, Poland
[6]Department of Rheumatology, Faculdade de Medicina e Centro Hospitalar, Universidade de Coimbra, Coimbra, Portugal
[7]Department of Surgery, University of Zaragoza, Zaragosa, Spain
[8]Department of Surgery, Division of Orthopaedics, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[9]Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
[10]Faculty of Medicine, St. Joseph University, Bellevue University Medical Center, Beirut, Lebanon
[11]University College London, London, UK
[12]Department of Orthopaedic Surgery, Hopital Rangueil, Centre Hospitalier Universitaire de Toulouse, Toulouse, France
[13]Past President EFORT, University Ulm, Germany
[14]Department of Internal Medicine III, National Institute of Rheumatology and Physiotherapy, Rheumatology Chair, Semmelweis University, Budapest, Hungary
[15]Department of Rheumatology, INSERM 1153, Cochin Hospital, Paris Descartes University, Paris, France
[16]Research Laboratory and Academic, Division of Clinical Rheumatology, Department of Internal Medicine, University of Genova, Genova, Italy
[17]Department of Orthopedics and Traumatology, Tor Vergata University of Rome, Rome, Italy
[18]Department of Family Medicine, Maastricht University, CAPHRI—School for Public Health and Primary Care, Maastricht, The Netherlands
[19]Bone and Joint Research Group, Knowledge Spa, Royal Cornwall Hospital, Truro, UK
[20]Department of Internal Medicine, NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, The Netherlands
[21]Department of Internal Medicine, VieCuri Medical Center, Venlo, The Netherlands
[22]Department of Internal Medicine, Rheumatology, Maastricht University Medical Center, Maastricht, The Netherlands
[23]University Hasselt, Hasselt, Belgium

Acknowledgements The set of recommendations were initiated by both the European League Against Rheumatism and the European Federation of National Associations of Orthopaedics and Traumatology board; we are very grateful for this support.

Contributors All authors were contributor to the design of the study including the formulation of research questions, to the analysis of and the discussion around the literature and have read and given comments on the manuscript.

Funding Two one-day meetings were organised, these were financially supported by unrestricted grants from the European League Against Rheumatism and the European Federation of National Associations of Orthopaedics and Traumatology.

Competing interests WFL reports personal fees (speakers fee/advisory boards) from Amgen, Eli Lilly, Novartis and Merck. KED reports personal fees from Agnovos, Amgen, Bayer, Bertelsmann, Heel, Janssen, Eli Lilly, Merck, Sanofi and UCB. HB-F reports to have been an invited speaker/on advisory boards by Roche Diagnostics, Nestlé, Pfizer, WILD, Sanofi and Sandoz. Investigator initiated funding from Nestlé, Pfizer, WILD and DSM Nutritional Products. EC reports remuneration from Amgen and Regeneron during the conduct of the study. TK reports personal fees from AbbVie, Biogen, BMS, Boehringer Ingelheim, Celltrion, Eli Lilly, Epirus, Janssen, Merck-Serono, MSD, Mundipharma, Novartis, Oktal, Orion Pharma, Hospira/Pfizer, Roche, Sandoz and from UCB Pharma. CR reports personal fees from Amgen, MSD, Eli Lilly and grants from Ultragenyx. PG reports grants and other from Amgen and Eli Lilly, and grants from Pfizer, MSD, UCB, Abbott, BMS, Novartis, Roche and Will Pharma.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Kanis JA, Borgström F, Compston J, *et al*. SCOPE: a scorecard for osteoporosis in Europe. *Arch Osteoporos* 2013;8:144.
2 Landewé RB, Günther KP,, Lukas C, *et al*. EULAR/EFORT recommendations for the diagnosis and initial management of patients with acute or recent onset swelling of the knee. *Ann Rheum Dis* 2010;69:12–19.
3 Dougados M, Betteridge N, Burmester GR, *et al*. EULAR standardised operating procedures for the elaboration, evaluation, dissemination, and implementation of recommendations endorsed by the EULAR standing committees. *Ann Rheum Dis* 2004;63:1172–6.
4 van der Heijde D, Aletaha D, Carmona L, *et al*. 2014 Update of the EULAR standardised operating procedures for EULAR-endorsed recommendations. *Ann Rheum Dis* 2015;74:8–13.
5 Tosounidis TH, Sheikh H, Stone MH, *et al*. Pain relief management following proximal femoral fractures: Options, issues and controversies. *Injury* 2015;46(Suppl 5):S52–8.
6 Abou-Setta AM, Beaupre LA, Rashiq S, *et al*. Comparative effectiveness of pain management interventions for hip fracture: a systematic review. *Ann Intern Med* 2011;155:234–45.
7 Wilson H. Multi-disciplinary care of the patient with acute hip fracture: How to optimise the care for the elderly, traumatised patient at and around the time of the fracture to ensure the best short-term outcome as a foundation for the best long-term outcome. *Best Pract Res Clin Rheumatol* 2013;27:717–30.
8 Roberts KC, Brox WT. From evidence to application: AAOS clinical practice guideline on management of hip fractures in the elderly. *J Orthop Trauma* 2015;29:119–20.
9 Ftouh S, Morga A, Swift C, *et al*. Management of hip fracture in adults: summary of NICE guidance. *BMJ* 2011;342:d3304.
10 Mears SC, Kates SL. A guide to improving the care of patients with fragility fractures, Edition 2. *Geriatr Orthop Surg Rehabil* 2015;6:58–120.
11 Martinez-Reig M, Ahmad L, Duque G. The orthogeriatrics model of care: systematic review of predictors of institutionalization and mortality in post-hip fracture patients and evidence for interventions. *J Am Med Dir Assoc* 2012;13:770–7.
12 Simunovic N, Devereaux PJ, Sprague S, *et al*. Effect of early surgery after hip fracture on mortality and complications: systematic review and meta-analysis. *CMAJ* 2010;182:1609–16.
13 Roberts KC, Brox WT. AAOS clinical practice guideline: management of hip fractures in the elderly. *J Am Acad Orthop Surg* 2015;23:138–40.
14 Menendez ME, Ring D. Does the timing of surgery for proximal humeral fracture affect inpatient outcomes? *J Shoulder Elbow Surg* 2014;23:1257–62.
15 Ellis G, Whitehead MA, Robinson D, *et al*. Comprehensive geriatric assessment for older adults admitted to hospital: meta-analysis of randomised controlled trials. *BMJ* 2011;343:d6553.
16 De Rui M, Veronese N, Manzato E, *et al*. Role of comprehensive geriatric assessment in the management of osteoporotic hip fracture in the elderly: an overview. *Disabil Rehabil* 2013;35:758–65.
17 Prestmo A, Hagen G, Sletvold O, *et al*. Comprehensive geriatric care for patients with hip fractures: a prospective, randomised, controlled trial. *Lancet* 2015;385:1623–33.
18 Grigoryan KV, Javedan H, Rudolph JL. Orthogeriatric care models and outcomes in hip fracture patients: a systematic review and meta-analysis. *J Orthop Trauma* 2014;28:e49–55.
19 Kammerlander C, Roth T, Friedman SM, *et al*. Ortho-geriatric service—a literature review comparing different models. *Osteoporos Int* 2010;21(Suppl 4):S637–46.

20 Sabharwal S, Wilson H. Orthogeriatrics in the management of frail older patients with a fragility fracture. *Osteoporos Int* 2015;26:2387–99.
21 Carpintero P, Caeiro JR, Carpintero R, *et al*. Complications of hip fractures: a review. *World J Orthop* 2014;5:402–11.
22 Bruce AJ, Ritchie CW, Blizard R, *et al*. The incidence of delirium associated with orthopedic surgery: a meta-analytic review. *Int Psychogeriatr* 2007;19:197–214.
23 Marsland D, Mears SC, Kates SL. Venous thromboembolic prophylaxis for hip fractures. *Osteoporos Int* 2010;21(Suppl 4):S593–604.
24 Geerts WH, Bergqvist D, Pineo GF, *et al*. Prevention of venous thromboembolism: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest* 2008;133(6 Suppl):381S–453S.
25 Handoll HH, Farrar MJ, McBirnie J, *et al*. Heparin, low molecular weight heparin and physical methods for preventing deep vein thrombosis and pulmonary embolism following surgery for hip fractures. *Cochrane Database Syst Rev* 2002(4): CD000305.
26 Comfort EH. Reducing pressure ulcer incidence through Braden Scale risk assessment and support surface use. *Adv Skin Wound Care* 2008;21:330–4.
27 Duncan DG BS, Hood K, Johansen A. Using dietetic assistants to improve the outcome of hip fracture: a randomised controlled trial of nutritional support in an acute trauma ward. *Age Ageing* 2006;35:148–53.
28 Bell JJ, Bauer JD, Capra S, *et al*. Multidisciplinary, multi-modal nutritional care in acute hip fracture inpatients—results of a pragmatic intervention. *Clin Nutr* 2014;33:1101–7.
29 Kammerlander C, Erhart S, Doshi H, *et al*. Principles of osteoporotic fracture treatment. *Best Pract Res Clin Rheumatol* 2013;27:757–69.
30 Wong TC, Chiu Y, Tsang WL, *et al*. Casting versus percutaneous pinning for extra-articular fractures of the distal radius in an elderly Chinese population: a prospective randomised controlled trial. *J Hand Surg Eur Vol* 2010;35:202–8.
31 Arora R, Lutz M, Deml C, *et al*. A prospective randomized trial comparing nonoperative treatment with volar locking plate fixation for displaced and unstable distal radial fractures in patients sixty-five years of age and older. *J Bone Joint Surg Am* 2011;93:2146–53.
32 Bartl C, Stengel D, Bruckner T, *et al*. The treatment of displaced intra-articular distal radius fractures in elderly patients. *Dtsch Arztebl Int* 2014;111: 779–87.
33 Bukata SV, Kates SL, O'Keefe RJ. Short-term and long-term orthopaedic issues in patients with fragility fractures. *Clin Orthop Relat Res* 2011;469:2225–36.
34 Longo UG, Loppini M, Denaro L, *et al*. Conservative management of patients with an osteoporotic vertebral fracture: a review of the literature. *J Bone Joint Surg Br* 2012;94:152–7.
35 Staples MP, Kallmes DF, Comstock BA, *et al*. Effectiveness of vertebroplasty using individual patient data from two randomised placebo controlled trials: meta-analysis. *BMJ* 2011;343:d3952.
36 Anderson PA, Froyshteter AB, Tontz WL Jr. Meta-analysis of vertebral augmentation compared with conservative treatment for osteoporotic spinal fractures. *J Bone Miner Res* 2013;28:372–82.
37 Papanastassiou ID, Phillips FM, Van Meirhaeghe J, *et al*. Comparing effects of kyphoplasty, vertebroplasty, and non-surgical management in a systematic review of randomized and non-randomized controlled studies. *Eur Spine J* 2012;21:1826–43.
38 Bauer DC. Vertebral augmentation vs nonsurgical therapy: improved symptoms, improved survival, or neither? *JAMA Intern Med* 2013;173:1522–3.
39 Florschutz AV, Langford JR, Haidukewych GJ, *et al*. Femoral neck fractures: current management. *J Orthop Trauma* 2015;29:121–9.
40 Burgers PT, Van Geene AR, Van den Bekerom MP, *et al*. Total hip arthroplasty versus hemiarthroplasty for displaced femoral neck fractures in the healthy elderly: a meta-analysis and systematic review of randomized trials. *Int Orthop* 2012;36:1549–60.
41 Yu L, Wang Y, Chen J. Total hip arthroplasty versus hemiarthroplasty for displaced femoral neck fractures: meta-analysis of randomized trials. *Clin Orthop Relat Res* 2012;470:2235–43.
42 van den Bekerom MP, Hilverdink EF, Sierevelt IN, *et al*. A comparison of hemiarthroplasty with total hip replacement for displaced intracapsular fracture of the femoral neck: a randomised controlled multicentre trial in patients aged 70 years and over. *J Bone Joint Surg Br* 2010;92:1422–8.
43 Jiang J, Yang CH, Lin Q, *et al*. Does arthroplasty provide better outcomes than internal fixation at mid- and long-term follow-up? A meta-analysis. *Clin Orthop Relat Res* 2015;473:2672–9.
44 Joglekar SB, Lindvall EM, Martirosian A. Contemporary management of subtrochanteric fractures. *Orthop Clin North Am* 2015;46:21–35.
45 Rangan A, Handoll H, Brealey S, *et al*. Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. *JAMA* 2015;313:1037–47.
46 Handoll HH, Brorson S. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev* 2015(11):CD000434.
47 Maier D, Jaeger M, Izadpanah K, *et al*. Proximal humeral fracture treatment in adults. *J Bone Joint Surg Am* 2014;96:251–61.
48 Marsh D, Akesson K, Beaton DE, *et al*. Coordinator-based systems for secondary prevention in fragility fracture patients. *Osteoporos Int* 2011;22:2051–65.
49 Ganda K, Puech M, Chen JS, *et al*. Models of care for the secondary prevention of osteoporotic fractures: a systematic review and meta-analysis. *Osteoporos Int* 2013;24:393–406.
50 McLellan AR, Gallacher SJ, Fraser M, *et al*. The fracture liaison service: success of a program for the evaluation and management of patients with osteoporotic fracture. *Osteoporos Int* 2003;14:1028–34.
51 Jaglal SB, Donescu OS, Bansod V, *et al*. Impact of a centralized osteoporosis coordinator on post-fracture osteoporosis management: a cluster randomized trial. *Osteoporos Int* 2012;23:87–95.
52 Majumdar SR, Beaupre LA, Harley CH, *et al*. Use of a case manager to improve osteoporosis treatment after hip fracture: results of a randomized controlled trial. *Arch Intern Med* 2007;167:2110–15.
53 Morrish DW, Beaupre LA, Bell NR, *et al*. Facilitated bone mineral density testing versus hospital-based case management to improve osteoporosis treatment for hip fracture patients: additional results from a randomized trial. *Arthritis Rheum* 2009;61:209–15.
54 Eisman JA, Bogoch ER, Dell R, *et al*. Making the first fracture the last fracture: ASBMR task force report on secondary fracture prevention. *J Bone Miner Res* 2012;27:2039–46.
55 Akesson K, Marsh D, Mitchell PJ, *et al*. Capture the fracture: a best practice framework and global campaign to break the fragility fracture cycle. *Osteoporos Int* 2013;24:2135–52.
56 Garvan Institut. Fracture Risk Calculator. Secondary Fracture Risc Calculator. http://www.garvan.org.au/promotions/bone-fracture-risk/calculator/
57 Clin Risk. Q-Fracture Risc Calculator. Secondary Q-Fracture Risc Calculator. http://www.qfracture.org/
58 Kanis JA, McCloskey EV, Johansson H, *et al*. Case finding for the management of osteoporosis with FRAX—assessment and intervention thresholds for the UK. *Osteoporos Int* 2008;19:1395–408.
59 Nguyen ND, Frost SA, Center JR, *et al*. Development of a nomogram for individualizing hip fracture risk in men and women. *Osteoporos Int* 2007;18:1109–17.
60 Hippisley-Cox J, Coupland C. Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. *BMJ* 2009;339:b4229.
61 Appelman-Dijkstra NM, Papapoulos SE. Prevention of incident fractures in patients with prevalent fragility fractures: current and future approaches. *Best Pract Res Clin Rheumatol* 2013;27:805–20.
62 Aspray TJ. Fragility fracture: recent developments in risk assessment. *Ther Adv Musculoskelet Dis* 2015;7:17–25.
63 Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ* 1996;312:1254–9.
64 Gallacher SJ, Gallagher AP, McQuillian C, *et al*. The prevalence of vertebral fracture amongst patients presenting with non-vertebral fractures. *Osteoporos Int* 2007;18:185–92.
65 Siris ES, Adler R, Bilezikian J, *et al*. The clinical diagnosis of osteoporosis: a position statement from The National Bone Health Alliance Working Group. *Osteoporos Int* 2014;25:1439–43.
66 Bours SP, van den Bergh JP, van Geel TA, *et al*. Secondary osteoporosis and metabolic bone disease in patients 50 years and older with osteoporosis or with a recent clinical fracture: a clinical perspective. *Curr Opin Rheumatol* 2014;26:430–9.
67 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;362:1225–30.
68 American Academy of Orthopaedic Surgeons. Management of Hip Fractures in the Elderly. Secondary Management of Hip Fractures in the Elderly. 2014. http://www.aaos.org/research/guidelines/HipFxGuideline.pdf
69 British Orthopaedic Association. The care of patients with fragility fracture. Secondary care of patients with fragility fracture. 2007. http://www.fractures.com/pdf/BOA-BGS-Blue-Book.pdf
70 Neuburger J, Currie C, Wakeman R, *et al*. The impact of a national clinician-led audit initiative on care and mortality after hip fracture in England: an external evaluation using time trends in non-audit data. *Med Care* 2015;53:686–91.
71 Ban I, Palm H, Birkelund L, *et al*. Implementing, adapting, and validating an evidence-based algorithm for hip fracture surgery. *J Orthop Trauma* 2014;28: e21–6.
72 Kammerlander C, Gosch M, Kammerlander-Knauer U, *et al*. Long-term functional outcome in geriatric hip fracture patients. *Arch Orthop Trauma Surg* 2011;131:1435–44.
73 Suhm N, Kaelin R, Studer P, *et al*. Orthogeriatric care pathway: a prospective survey of impact on length of stay, mortality and institutionalisation. *Arch Orthop Trauma Surg* 2014;134:1261–9.
74 Handoll HH, Cameron ID, Mak JC, *et al*. Multidisciplinary rehabilitation for older people with hip fractures. *Cochrane Database Syst Rev* 2009;(4):CD007125.
75 Halbert J, Crotty M, Whitehead C, *et al*. Multi-disciplinary rehabilitation after hip fracture is associated with improved outcome: a systematic review. *J Rehabil Med* 2007;39:507–12.

76 Swanson CE, Day GA, Yelland CE, *et al*. The management of elderly patients with femoral fractures. A randomised controlled trial of early intervention versus standard care. *Med J Aust* 1998;169:515–18.
77 Crotty M, Whitehead CH, Gray S, *et al*. Early discharge and home rehabilitation after hip fracture achieves functional improvements: a randomized controlled trial [with consumer summary]. *Clin Rehabil* 2002;16:406–13.
78 Darowski A. *The Care of Patients with Fragility Fracture ("Blue Book")*. British Orthopaedic Association, 2007.
79 Oldmeadow LB, Edwards ER, Kimmel LA, *et al*. No rest for the wounded: early ambulation after hip surgery accelerates recovery. *ANZ J Surg* 2006;76:607–11.
80 Baker PA, Evans OM, Lee C, *et al*. Treadmill gait retraining following fractured neck-of-femur. *Arch Phys Med Rehabil* 1991;72:649–52.
81 Mendelsohn ME, Overend TJ, Connelly DM, *et al*. Improvement in aerobic fitness during rehabilitation after hip fracture. *Arch Phys Med Rehabil* 2008;89:609–17.
82 Giangregorio LM, Macintyre NJ, Thabane L, *et al*. Exercise for improving outcomes after osteoporotic vertebral fracture. *Cochrane Database Syst Rev* 2013;1:CD008618.
83 Body JJ, Bergmann P, Boonen S, *et al*. Non-pharmacological management of osteoporosis: a consensus of the Belgian Bone Club. *Osteoporos Int* 2011;22:2769–88.
84 Siris ES, Gehlbach S, Adachi JD, *et al*. Failure to perceive increased risk of fracture in women 55 years and older: the Global Longitudinal Study of Osteoporosis in Women (GLOW). *Osteoporos Int* 2011;22:27–35.
85 Little EA, Eccles MP. A systematic review of the effectiveness of interventions to improve post-fracture investigation and management of patients at risk of osteoporosis. *Implement Sci* 2010;5:80.
86 Sale JE, Beaton D, Posen J, *et al*. Systematic review on interventions to improve osteoporosis investigation and treatment in fragility fracture patients. *Osteoporos Int* 2011;22:2067–82.
87 Laliberte MC, Perreault S, Jouini G, *et al*. Effectiveness of interventions to improve the detection and treatment of osteoporosis in primary care settings: a systematic review and meta-analysis. *Osteoporos Int* 2011;22:2743–68.
88 O'Brien LK, Armstrong AD, Hassenbein SE, *et al*. Evaluation of patients' response toward osteoporosis letter intervention versus phone call plus letter intervention. *Geriatr Orthop Surg Rehabil* 2015;6:246–50.
89 Ryan R, Santesso N, Lowe D, *et al*. Interventions to improve safe and effective medicines use by consumers: an overview of systematic reviews. *Cochrane Database Syst Rev* 2014;4:CD007768.
90 Reid IR, Bristow SM, Bolland MJ. Calcium supplements: benefits and risks. *J Intern Med* 2015;278:354–68.
91 Sprague S, Petrisor B, Scott T, *et al*. What is the role of vitamin D supplementation in acute fracture patients? A systematic review and meta-analysis of the prevalence of hypovitaminosis D and supplementation efficacy. *J Orthop Trauma* 2016;30:53–63.
92 Bischoff-Ferrari HA, Dawson-Hughes B, Staehelin HB, *et al*. Fall prevention with supplemental and active forms of vitamin D: a meta-analysis of randomised controlled trials. *BMJ* 2009;339:b3692.
93 Bischoff-Ferrari HA, Willett WC, Orav EJ, *et al*. A pooled analysis of vitamin D dose requirements for fracture prevention. *N Engl J Med* 2012;367:40–9.
94 Bischoff-Ferrari HA, Willett WC, Wong JB, *et al*. Prevention of nonvertebral fractures with oral vitamin D and dose dependency: a meta-analysis of randomized controlled trials. *Arch Intern Med* 2009;169:551–61.
95 Bruyere O, Cavalier E, Souberbielle JC, *et al*. Effects of vitamin D in the elderly population: current status and perspectives. *Arch Public Health* 2014;72:32.
96 Sanders KM, Stuart AL, Williamson EJ, *et al*. Annual high-dose oral vitamin D and falls and fractures in older women: a randomized controlled trial. *JAMA* 2010;303:1815–22.
97 Bischoff-Ferrari HA, Dawson-Hughes B, Orav EJ, *et al*. Monthly high-dose vitamin D treatment for the prevention of functional decline: a randomized clinical trial. *JAMA Intern Med* 2016;176:175–83.
98 Lyles KW, Colon-Emeric CS, Magaziner JS, *et al*. Zoledronic acid and clinical fractures and mortality after hip fracture. *N Engl J Med* 2007;357:1799–809.
99 Black DM, Cummings SR, Karpf DB, *et al*. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. *Lancet* 1996;348:1535–41.
100 Black DM, Delmas PD, Eastell R, *et al*. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. *N Engl J Med* 2007;356:1809–22.
101 Cummings SR, San Martin J, McClung MR, *et al*. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. *N Engl J Med* 2009;361:756–65.
102 McClung MR, Geusens P, Miller PD, *et al*. Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. *N Engl J Med* 2001;344:333–40.
103 Neer RM, Arnaud CD, Zanchetta JR, *et al*. Effect of parathyroid hormone (1–34) on fractures and bone mineral density in postmenopausal women with osteoporosis. *N Engl J Med* 2001;344:1434–41.
104 Geusens PP, van den Bergh JP. Bone: new guidelines for multistep fracture prevention in men. *Nat Rev Rheumatol* 2012;8:568–70.
105 Fried TR. Shared decision making—finding the sweet spot. *N Engl J Med* 2016;374:104–6.
106 Lewiecki EM. Risk communication and shared decision making in the care of patients with osteoporosis. *J Clin Densitom* 2010;13:335–45.
107 Eekman DA, van Helden SH, Huisman AM, *et al*. Optimizing fracture prevention: the fracture liaison service, an observational study. *Osteoporos Int* 2014;25:701–9.

OPEN ACCESS

EXTENDED REPORT

Lesinurad in combination with allopurinol: a randomised, double-blind, placebo-controlled study in patients with gout with inadequate response to standard of care (the multinational CLEAR 2 study)

Thomas Bardin,[1] Robert T Keenan,[2] Puja P Khanna,[3] Jeff Kopicko,[4] Maple Fung,[5] Nihar Bhakta,[5] Scott Adler,[6] Chris Storgard,[5] Scott Baumgartner,[7] Alexander So[8]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-209213).

For numbered affiliations see end of article.

Correspondence to
Professor Thomas Bardin, Rhumatologie, Hôpital Lariboisière, Assistance Publique Hôpitaux de Paris, 2, rue A Paré, 75010 Paris; thomas.bardin@aphp.fr

JK and CS are former employees.

Received 17 January 2016
Revised 25 August 2016
Accepted 5 October 2016
Published Online First 7 November 2016

ABSTRACT

Objectives Determine the efficacy and safety of daily lesinurad (200 or 400 mg orally) added to allopurinol in patients with serum uric acid (sUA) above target in a 12-month, randomised, phase III trial.

Methods Patients on allopurinol ≥300 mg (≥200 mg in moderate renal impairment) had sUA level of ≥6.5 mg/dL (≥387 μmol/L) at screening and two or more gout flares in the prior year. Primary end point was the proportion of patients achieving sUA level of <6.0 mg/dL (<357 μmol/L) (month 6). Key secondary end points were mean gout flare rate requiring treatment (months 7 through 12) and proportions of patients with complete resolution of one or more target tophi (month 12). Safety assessments included adverse events and laboratory data.

Results Patients (n=610) were predominantly male, with mean (±SD) age 51.2±10.90 years, gout duration 11.5±9.26 years and baseline sUA of 6.9±1.2 mg/dL (410±71 μmol/L). Lesinurad at 200 and 400 mg doses, added to allopurinol, significantly increased proportions of patients achieving sUA target versus allopurinol-alone therapy by month 6 (55.4%, 66.5% and 23.3%, respectively, p<0.0001 both lesinurad+allopurinol groups). In key secondary end points, there were no statistically significant treatment-group differences favouring lesinurad. Lesinurad was generally well tolerated; the 200 mg dose had a safety profile comparable with allopurinol-alone therapy. Renal-related adverse events occurred in 5.9% of lesinurad 200 mg +allopurinol, 15.0% of lesinurad 400 mg+allopurinol and 4.9% of allopurinol-alone groups, with serum creatinine elevation of ≥1.5× baseline in 5.9%, 15.0% and 3.4%, respectively. Serious treatment-emergent adverse events occurred in 4.4% of lesinurad 200 mg +allopurinol, in 9.5% of lesinurad 400 mg+allopurinol and in 3.9% of allopurinol-alone groups, respectively.

Conclusion Lesinurad added to allopurinol demonstrated superior sUA lowering versus allopurinol-alone therapy and lesinurad 200 mg was generally well tolerated in patients with gout warranting additional therapy.

Trial registration number NCT01493531.

Linked

► http://dx.doi.org/10.1136/annrheumdis-2016-210519

CrossMark

To cite: Bardin T, Keenan RT, Khanna PP, *et al*. *Ann Rheum Dis* 2017;**76**:811–820.

INTRODUCTION

Gout is an inflammatory arthritis characterised by the deposition of monosodium urate (MSU) crystals in the joints, tendons and other connective tissues. Crystal deposition secondary to long-standing hyperuricemia can be reversed by lowering the concentration of serum uric acid (sUA) below the MSU saturation point—leading, in the long term, to the potential disappearance of signs and symptoms of gout. As a result, current management guidelines recommend maintenance of sUA to <6.0 mg/dL (<357 μmol/L) in patients with gout.[1–3]

Allopurinol is recommended as a first-line urate-lowering therapy (ULT).[2 4] However, clinical trials have demonstrated that >50% of patients do not achieve sustained reductions in sUA at the most commonly used allopurinol dose of 300 mg.[5–8] Lesinurad (RDEA594) is a novel, selective uric acid reabsorption inhibitor (SURI) for treatment of gout in combination with xanthine oxidase inhibitors. Lesinurad inhibits URAT1, a uric acid transporter responsible for the reabsorption of uric acid from the renal tubular lumen.[9–11] Lesinurad in combination with allopurinol therefore provides a dual mechanism for sUA lowering—an increase in excretion of uric acid and a reduction in urate production.

Clinical studies have demonstrated that lesinurad in combination with allopurinol reduces mean sUA concentrations and increases proportions of patients who achieve sUA targets.[12–14] The current phase III study—Combining Lesinurad with Allopurinol Standard of Care in Inadequate Responders (CLEAR 2)—is one of two replicate, randomised, double-blind, placebo-controlled, multicentre studies to investigate lesinurad in combination with allopurinol in patients with gout. CLEAR 1 was performed within the USA, included 603 patients with gout and provided outcomes similar to the CLEAR 2 study.[15]

METHODS

Study design

CLEAR 2 was an international, phase III trial to investigate the efficacy and safety of two lesinurad doses (200 or 400 mg oral, once daily) in combination with allopurinol, versus allopurinol combined with placebo (the control arm), in patients demonstrating inadequate response to standard-of-care allopurinol (ClinicalTrials.gov Identifier: NCT01493531). The study was conducted in 12 countries in Europe, North America, South Africa, Australia and New Zealand between December 2011 and July 2014.

CLEAR 2 included a screening period of approximately 28 days, including a run-in of approximately 14 days on gout flare prophylaxis and 12-month double-blind treatment (figure 1). The study was conducted in accordance with Independent Ethics Committee E6 Good Clinical Practice, the Declaration of Helsinki (October 2008) and all applicable local regulatory requirements.

Patients

Male or female patients aged 18–85 years with a diagnosis of gout, body mass index <45 kg/m^2, inadequate hypouricaemic response to standard-of-care allopurinol and two or more gout flares in the previous 12 months were eligible for study inclusion. Patients were included if they met the 1977 American Rheumatism Association preliminary classification criteria for gout.[16] Patients were required to have received allopurinol as the sole ULT for ≥8 weeks prior to screening at a dose assessed medically appropriate by the treating physician (minimum 300 mg/day (200 mg in moderate renal impairment)[17] up to 800 or 900 mg, depending on locally approved dose). sUA was required to be ≥6.5 mg/dL (≥387 µmol/L) at screening and ≥6.0 mg/dL (≥357 µmol/L) approximately 7 days prior to start of treatment on day 1.

Patients with estimated creatinine clearance (eCrCl) <30 mL/min were excluded from study. Patients with a history of kidney stones were permitted. Complete exclusion criteria are included in the online supplementary material 1.

Study medications

Eligible patients were randomised by double-blind method to one of three treatment groups (lesinurad 200 mg, lesinurad 400 mg or placebo) in 1:1:1 ratio, added to continued treatment with allopurinol at pre-study dose. Randomisation at study sites used a centralised Interactive Voice Response System/Interactive Web Response System.

Doses of lesinurad or matching placebo were taken once daily in the morning with food and one cup of water. Compliance was assessed from dispensing records and verification of returned medication packaging. Concomitant medication use was recorded at each study visit.

Gout flare prophylaxis was initiated at day −14, that is, the same time as sponsor-provided allopurinol. Prophylaxis consisted of colchicine (0.5 or 0.6 mg/day, as locally available) or a non-steroidal anti-inflammatory drug (NSAID, dosed according to local prescribing practice, with or without proton-pump inhibitor) for patients who were intolerant to or had contraindications to colchicine. Gout flare prophylaxis was continued through month 5, unless patients became intolerant or developed toxicity to prophylaxis.

Patients were encouraged to drink 2 L of fluid a day and remain well hydrated, following American College of Rheumatology guidelines for management of gout.[2]

Assessments

Efficacy assessments

The primary efficacy end point was the proportion of patients in each treatment group with sUA <6.0 mg/dL (<357 µmol/L) by month 6. Other sUA-related end points included proportions of patients with sUA <6.0 mg/dL (<357 µmol/L), <5.0 mg/dL (<297 µmol/L) and <4.0 mg/dL (<238 µmol/L) and mean absolute and mean percentage changes from baseline in sUA at each visit.

Two key secondary end points included: (1) mean rate of gout flares requiring treatment for the 6-month period from end of month 6 to end of month 12, reported on a daily electronic patient diary. This key secondary end point included only

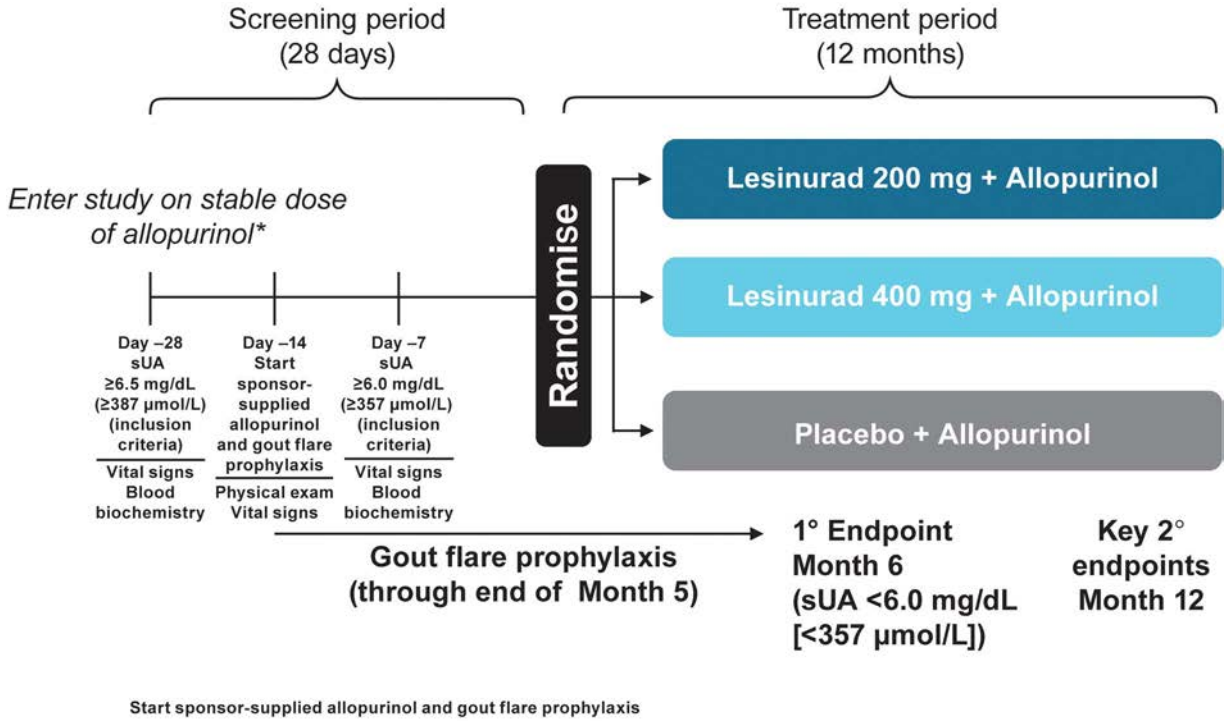


Figure 1 CLEAR 2 trial design is shown. *200 mg permitted for renally impaired. Maximum allopurinol dose: 800 or 900 mg, according to local label. Randomisation was stratified at day −7 by renal function (ie, estimated eCrCl ≥60 vs <60 mL/min, calculated by the Cockcroft-Gault formula using ideal body weight) and by tophus status during screening (ie, one or more tophus versus no tophi). eCrCl, estimated creatinine clearance; sUA, serum uric acid.

clinically relevant gout flares, which were those requiring either an increase in current medication or new medication and (2) proportion of patients with target tophi at baseline who experienced complete resolution of one or more target tophi by month 12, that is, 100% decrease in tophus area. Target tophi (up to five per patient) were tophi on the hands/wrists and/or feet/ankles measured by digital Vernier callipers at ≥5 and ≤20 mm in longest diameter.[18] Permitted treatments for gout flares were colchicine, analgesics and/or anti-inflammatory medications, including oral and intra-articular corticosteroids.

Safety assessments

Safety assessments included treatment-emergent adverse events (TEAEs; coded by Medical Dictionary for Regulatory Activities (V.14.0)), clinical laboratory data, physical examination, ECG and vital signs. Adverse events (AEs) of special interest included renal and cardiovascular (CV) safety assessments.

Assessments of renal safety included renal-related and kidney stone TEAEs (see online supplementary material 2) and clinical laboratory data, including serum creatinine (sCr), creatine kinase, urine protein-to-creatinine ratio and eCrCl levels. CV safety was of special interest because of the known high rates of CV risk factors in patients with gout.[19 20] An independent Cardiovascular Events Adjudication Committee (CEAC) routinely assessed AEs for potential CV relationship, with categorisation into major adverse CV events (MACEs) and non-MACE end points (see online supplementary material 3).[21]

Statistical analyses

Comparisons of response proportions based on sUA level between each lesinurad plus allopurinol group and the allopurinol-alone group were performed using the Cochran-Mantel-Haenszel (CMH) test statistic, stratified by day −7 renal function and tophus status during screening. A Bonferroni correction was used for the primary end point for each of the two treatment comparisons with allopurinol-alone therapy at an α level of 0.025. Testing of the key secondary end points hierarchically at an α level of 0.05 was gated on both dose contrasts being statistically significant for the primary end point. If only one of the primary end point dose contrasts was significant, then $\alpha=0.025$ for each key secondary end point within the surviving dose. All other efficacy end points were evaluated at $\alpha=0.05$ (nominal p value), two-sided, without multiplicity adjustment. Results for the primary end point of sUA response are expressed as proportions and p values. Patients with missing values at month 6 or month 12 for any reason were considered non-responders (non-responder imputation, NRI).[22] Key secondary end points were analysed using negative binomial regression (gout flares) or CMH test (tophus response). Mean rates of gout flares were adjusted for day −7 renal function, tophus status at screening and length of exposure to randomised study medication. The time points and analytical methods used in the study were agreed with multiple regulatory agencies.

Safety data are listed by treatment arm and are not subjected to statistical testing. TEAEs are coded by system organ class and preferred term and are listed according to incidence, severity, relation to study medication and relation to discontinuation. To better identify potential clinically relevant changes in sCr related to lesinurad by minimising discrepancies due to intra-subject variability, baseline sCr was defined as the highest value within 14 days prior to first dose of study medication. Relative increase in sCr (ie, ≥1.5× and ≥2.0× the baseline level at any time) was selected as the most clinically relevant sCr assessment.[23 24] Resolution of sCr elevation was defined as an sCr value returned to ≤1.2× baseline.

A sample size of approximately 600 patients was planned to be recruited, for an allocation of approximately 200 patients to each treatment arm. This sample size was calculated to provide greater than 90% power to detect a difference in response rate between treatment groups if the allopurinol-alone group had a 30% response rate and the lesinurad groups had response rates as low as 48% using Fisher's exact test, adjusting for multiplicity with $\alpha=0.025$, two-sided, for each test.

All randomised patients who received at least one dose of study medication were included in the intent-to-treat (ITT) population, which was the primary population for efficacy and safety assessments.

RESULTS

Patient disposition

Of the 2199 patients screened, 611 were randomised at 152 sites. Of the 611 randomised patients, 610 received at least one dose of study medication (figure 2).

Demographic characteristics and clinical history

Demographics and baseline disease characteristics were similar between treatment groups (table 1). Patients generally had long-standing symptomatic gout (mean (±SD) time since diagnosis 11.5±9.3 years) and elevated baseline sUA (mean 6.9±1.2 mg/dL (410±71 μmol/L)), with high rates of one or more predefined comorbidities (ie, CV risk factors or kidney stones) at 79.2%.

Most patients (84.1%) received allopurinol at a daily dose of 300 mg, with 6.6% receiving <300 mg and 9.3% receiving >300 mg; the overall dose range was 200–900 mg.

Study medications

Proportions of patients exhibiting ≥80% compliance with study medication were 97.6%, 94.1% and 94.5% in the allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively.

Efficacy assessments

Primary end point of sUA response and secondary sUA end points

Proportions of patients achieving sUA level of <6.0 mg/dL (<357 μmol/L) by month 6 (the primary end point) were 23.3%, 55.4% and 66.5% in the allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively, using NRI—significant differences were identified for both lesinurad+allopurinol groups versus allopurinol-alone (p<0.0001; CMH test) (figure 3).

Subgroup analyses based on age, sex, race, baseline sUA, comorbidities, renal function and thiazide diuretic use provided results consistent with primary analysis of the ITT population (see online supplementary material 4 for renal function and diuretic analyses).

Proportions of patients achieving the sUA target of <6.0 mg/dL (<357 μmol/L) were greater in the lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol versus allopurinol-alone group at all monthly assessments from month 1 to month 12 (nominal p<0.0001, all comparisons). Proportions of patients achieving sUA level of <5.0 mg/dL (<297 μmol/L) and <4.0 mg/dL (<238 μmol/L) were also greater in both lesinurad+allopurinol groups versus allopurinol-alone group at each monthly visit (sUA <5.0 mg/dL (<297 μmol/L): nominal p<0.0001, both comparisons; sUA <4.0 mg/dL (<238 μmol/L): nominal p<0.0001, both comparisons, except p<0.01 at

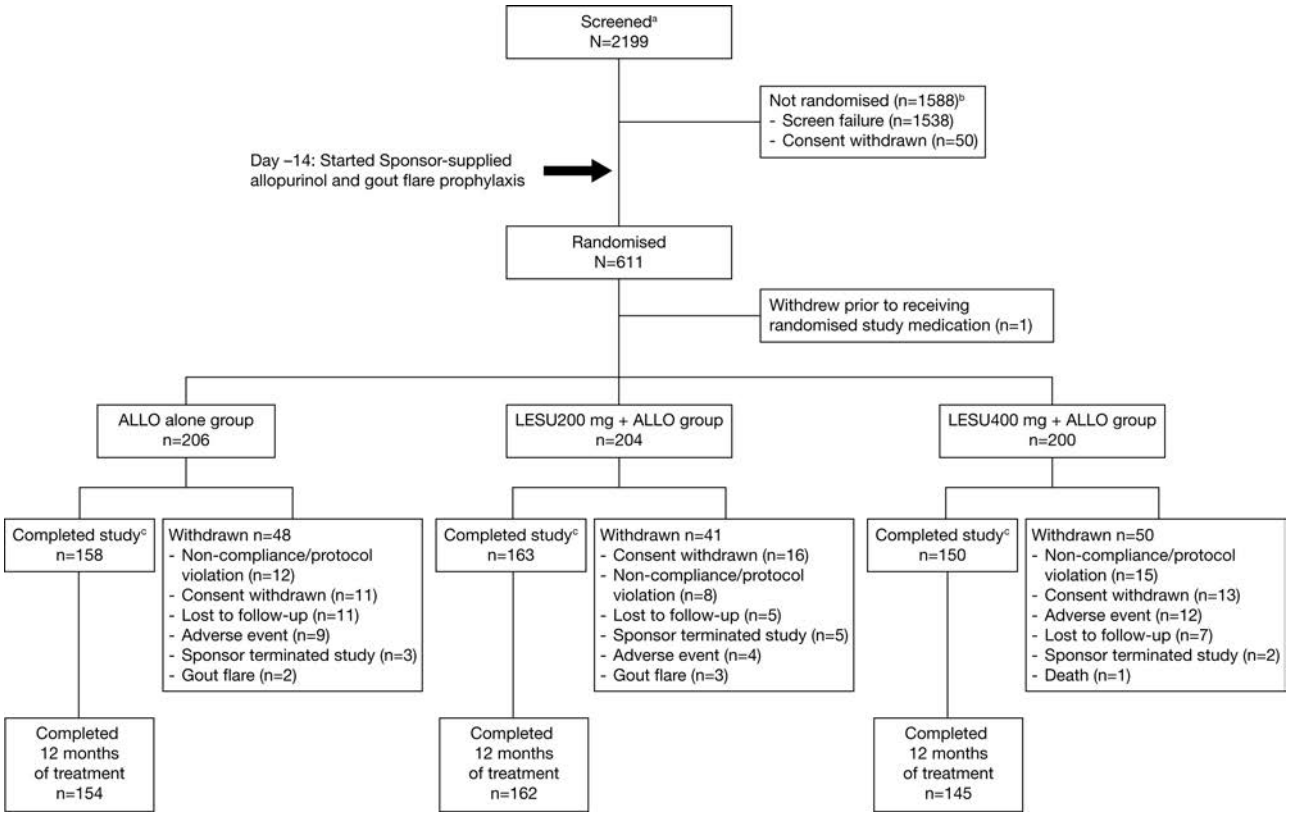


Figure 2 Patient disposition is shown. [a]Screened was defined as signing an informed consent form; [b]2 deaths reported for non-randomised patients during screening and [c]completed the study with or without completing randomised study medication. One additional death occurred in the LESU 400 mg+ALLO group. The subject experienced a serious adverse event and withdrew from the study. The primary reason for study withdrawal was reported as 'adverse event'. Of the 1538 screen failures, 1183 were related to inclusion criteria, 252 to exclusion criteria, 94 to both inclusion and exclusion criteria and 9 to other. ALLO, allopurinol; LESU, lesinurad.

month 1, lesinurad 200 mg+allopurinol). Figure 3 shows proportions of patients at each sUA threshold by months 6 and 12.

Mean sUA levels were lower in both lesinurad+allopurinol groups versus allopurinol-alone group at all time points (nominal $p<0.001$, both comparisons compared with allopurinol-alone therapy) (figure 4).

Secondary end point: gout flares requiring treatment

The gout flare rate and the proportions of patients with gout flares requiring treatment were low and similar in all groups throughout the study. Mean (±SE) rates of gout flares requiring treatment from the end of month 6 to end of month 12 were 0.83±0.13 for allopurinol-alone group versus 0.73±0.12 and 0.77±0.13, respectively, in the lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups ($p=0.57$ and 0.75 vs allopurinol-alone group). Proportions of patients with a gout flare requiring treatment through the study are shown in online supplementary figure S2.

Secondary end point: tophus resolution

The numbers of patients with one or more target tophi at baseline were low: 33, 35 and 29 in the allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively. In these respective groups, 33.3%, 31.4% and 27.6% of patients achieved complete resolution of one or more target tophi by month 12 ($p>0.05$, both lesinurad+allopurinol groups vs allopurinol-alone group).

Safety assessments

Adverse events

TEAEs were reported in 70.9%, 74.5% and 80.5% of the allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively (table 2). The majority of TEAEs in each group had a maximum severity of grade 1 or grade 2, based on Rheumatology Common Toxicity Criteria.[25] The most common individual TEAEs—reported for allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively—were upper respiratory tract infection (10.2%, 6.9%, 15.0%), hypertension (4.9%, 8.3%, 8.0%), arthralgia (4.4%, 11.8%, 3.0%), increased blood creatinine (3.4%, 3.9%, 9.5%) and diarrhoea (3.4%, 4.9%, 7.0%). The most common grade 3 or grade 4 TEAEs in these respective groups were increased blood creatine kinase (1.5%, 0.5%, 1.5%) and myocardial infarction (MI) (0%, 0%, 1.5%).

Serious TEAEs were reported in 3.9%, 4.4% and 9.5% of patients, respectively (table 2). Two deaths occurred in the lesinurad 400 mg+allopurinol group (pulmonary oedema and gastric cancer, respectively). TEAEs led to study-medication discontinuation in 5.3%, 3.4% and 9.5% of the allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively; the most common TEAE leading to discontinuation was increased blood creatinine (1.0%, 0% and 2.5%, respectively).

Renal safety analyses

Renal-related TEAEs occurred in 4.9%, 5.9% and 15.0%, of allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad

Table 1 Demographic and baseline disease characteristics (intent-to-treat population)

	ALLO alone (n=206)	Lesinurad 200 mg+ALLO (n=204)	Lesinurad 400 mg+ALLO (n=200)	Total (n=610)
Sex (n (%))				
Male	196 (95.1)	197 (96.6)	194 (97.0)	587 (96.2)
Female	10 (4.9)	7 (3.4)	6 (3.0)	23 (3.8)
Race (n (%))				
White	155 (75.2)	167 (81.9)	160 (80.0)	482 (79.0)
Black or African-American	22 (10.7)	15 (7.4)	21 (10.5)	58 (9.5)
Asian	14 (6.8)	10 (4.9)	9 (4.5)	33 (5.4)
Native Hawaiian or other Pacific Islander	5 (2.4)	3 (1.5)	2 (1.0)	10 (1.6)
American-Indian or Alaska Native	1 (0.5)	1 (0.5)	0	2 (0.3)
Other	8 (3.9)	4 (2.0)	6 (3.0)	18 (3.0)
Missing	0	0	1 (0.5)	1 (0.2)
Age (years)				
Mean (SD)	51.4 (10.56)	51.0 (11.11)	51.3 (11.08)	51.2 (10.90)
Min, max	21, 80	21, 82	18, 80	18, 82
BMI (kg/m^2)				
Mean (SD)	33.87 (6.19)	34.67 (6.43)	33.81 (6.68)	34.12 (6.44)
Min, max	21.91, 56.27	22.55, 55.63	22.76, 69.36	21.91, 69.36
Duration since gout diagnosis (years)				
Mean (SD)	11.31 (9.38)	12.25 (9.75)	11.02 (8.59)	11.53 (9.26)
Min, max	0.2, 53.0	0.5, 45.0	0.0, 47.4	0.0, 53.0
Presence of tophi at screening (n (%))				
Yes	48 (23.3)	49 (24.0)	47 (23.5)	144 (23.6)
No	158 (76.7)	155 (76.0)	153 (76.5)	466 (76.4)
Presence of ≥1 target tophus at baseline (n (%))				
Yes	33 (16.0)	35 (17.2)	29 (14.5)	97 (15.9)
No	173 (84.0)	169 (82.8)	171 (85.5)	513 (84.1)
No. of target tophi at baseline				
n	33	35	29	97
Mean (SD)	2.2 (1.36)	2.0 (1.34)	2.5 (1.53)	2.2 (1.40)
Min, max	1, 5	1, 5	1, 5	1, 5
No. of gout flares in the past 12 months				
Mean (SD)	5.8 (4.92)	6.7 (7.01)	6.1 (5.65)	6.2 (5.93)
Min, max	2, 30	2, 50	2, 48	2, 50
Renal function at baseline (mL/min) (n (%))				
eCrCl ≥90	72 (35.0)	80 (39.2)	85 (42.5)	237 (38.9)
eCrCl <90	133 (64.6)	124 (60.8)	114 (57.0)	371 (60.8)
eCrCl ≥60	165 (80.1)	175 (85.8)	170 (85.0)	510 (83.6)
eCrCl <60	40 (19.4)	29 (14.2)	29 (14.5)	98 (16.1)
CV risk factors (n (%))				
Hypertension	141 (68.4)	131 (64.2)	121 (60.5)	393 (64.4)
Hyperlipidaemia	76 (36.9)	86 (42.2)	93 (46.5)	255 (41.8)
Type 2 diabetes	28 (13.6)	31 (15.2)	26 (13.0)	85 (13.9)
History of kidney stones (n (%))	28 (13.6)	23 (11.3)	18 (9.0)	69 (11.3)
Baseline thiazide/thiazide-like diuretic use (n (%))	37 (18.0)	43 (21.1)	35 (17.5)	115 (18.9)
sUA at baseline (mg/dL) (μmol/L)				
Mean (SD)	7.0 (1.3) (416 (75))	6.8 (1.1) (407 (66))	6.9 (1.2) (410 (71))	6.9 (1.2) (410 (71))
Min, max	3.4, 11.3 (202, 672)	4.0, 11.3 (238, 672)	3.8, 11.0 (226, 654)	3.4, 11.3 (202, 672)
sUA category at baseline (n (%))				
<8.0 mg/dL (<476 μmol/L)	162 (78.6)	177 (86.8)	164 (82.0)	503 (82.5)
≥8.0 mg/dL (≥476 μmol/L)	44 (21.4)	27 (13.2)	36 (18.0)	107 (17.5)

Continued

Table 1 Continued

	ALLO alone (n=206)	Lesinurad 200 mg+ALLO (n=204)	Lesinurad 400 mg+ALLO (n=200)	Total (n=610)
Type of gout flare prophylaxis at baseline (n (%))				
Colchicine	159 (77.2)	181 (88.7)	167 (83.5)	507 (83.1)
NSAID	51 (24.8)	23 (11.3)	36 (18.0)	110 (18.0)
Both	8 (3.9)	4 (2.0)	3 (1.5)	15 (2.5)
Other or missing	4 (1.9)	4 (2.0)	0	8 (1.3)
Allopurinol dose at baseline (mg/day)				
Mean (SD)	308.7 (69.29)	313.5 (78.33)	314.8 (77.62)	312.3 (75.08)
Min, max	200, 600	200, 900	200, 900	200, 900

ALLO, allopurinol; BMI, body mass index; CV, cardiovascular; eCrCl, estimated creatinine clearance; NSAID, non-steroidal anti-inflammatory drug; sUA, serum uric acid.

400 mg+allopurinol groups, respectively. The most common renal-related TEAEs in these respective groups were increased blood creatinine (3.4%, 3.9%, 9.5%), increased blood urea (0%, 2.0%, 1.5%) and renal failure (0.5%, 1.0%, 1.5%). One patient (0.5%) in the allopurinol-alone group experienced a serious renal-related TEAE, versus no patients in the lesinurad 200 mg+allopurinol and two patients (1.0%) in the lesinurad 400 mg+allopurinol group. Kidney stone TEAEs were reported in 0.5%, 0% and 3.0%, respectively.

sCr elevation ≥1.5× baseline occurred in 3.4% (n=7), 5.9% (n=12) and 15.0% (n=30) of allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively. sCr elevation ≥1.5× was transient and reversible in most cases and the majority of sCr elevations resolved by the time of the next assessment; there were three unresolved sCr elevations in the allopurinol-alone group at last visit versus none in the lesinurad 200 mg+allopurinol and seven in the lesinurad 400 mg+allopurinol group (see online supplementary table S1). sCr elevation ≥2.0× baseline occurred in 0%, 2.0% (n=4) and 8.0% (n=16) of patients, respectively. Again, most elevations ≥2.0× baseline were transient and reversible; no sCr elevations ≥2.0× were unresolved at last visit in the lesinurad 200 mg +allopurinol group and five cases were unresolved in the lesinurad 400 mg+allopurinol group. In approximately two-thirds of sCr elevations, resolution occurred while patients continued on study medication.

In all treatment groups, proportions of patients with an sCr elevation ≥1.5× baseline tended to be higher for patients (1) who were taking an NSAID than colchicine; (2) who did not achieve target sUA at month 6 versus responders and (3) who had one or more tophi at screening versus those without tophi, although small subgroup sizes render interpretation difficult (see online supplementary table S2). There was no apparent association between sCr elevation and baseline renal function or other concomitant medications.

Renal function remained stable across the treatment groups, as measured by mean (SD) changes in eCrCl, from baseline to last value. Mean (±SD) changes in eCrCl in the allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups were 3.0±9.7, −0.5±11.5 and −5.7±13.9 mg/dL, respectively, from baseline to last value on treatment and were 1.8±11.7, 2.7±10.0 and 1.1±24.2 mg/dL from baseline to last value off treatment at follow-up (in patients not entering a separate extension study, n=133).

CV safety analyses

TEAEs were adjudicated as CV events in 5.3% (n=11 patients), 3.9% (n=8 patients) and 3.0% (n=6 patients) of allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively. CEAC-adjudicated criteria for MACE were met by three patients (four events, including three MIs and one death due to pulmonary oedema), all in

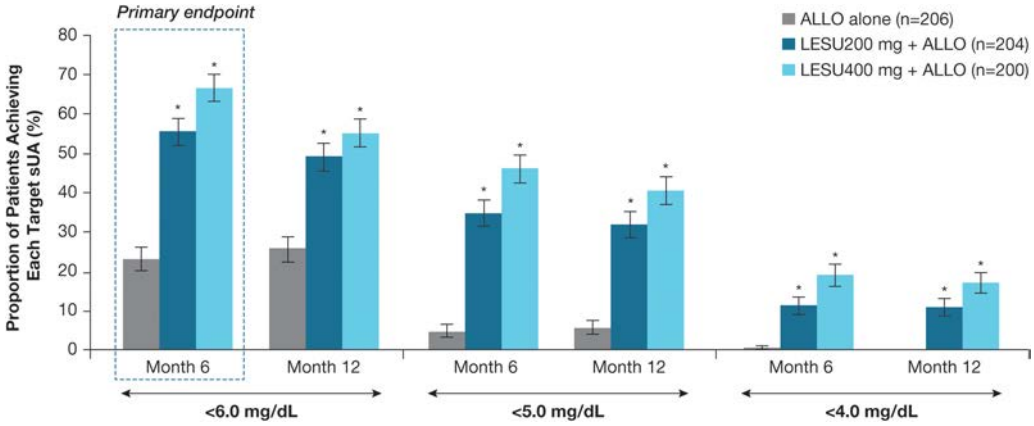


Figure 3 Proportions of patients achieving sUA target of <6.0 mg/dL (<357 μmol/L), <5.0 mg/dL (<297 μmol/L) and <4.0 mg/dL (<238 μmol/L), by months 6 and 12 (ITT population) are shown. Primary end point: proportion of patients achieving sUA target of <6.0 mg/dL (<357 μmol/L) by month 6. *p<0.0001. Note: Subjects missing sUA results were treated as non-responders. All comparisons used a two-sided Cochran-Mantel-Haenszel test stratified by day −7 renal function and tophus status during screening (randomised stratification factor values), with non-responder imputation and adjustment for multiple comparisons for the primary end point (Bonferroni correction). ALLO, allopurinol; ITT, intention to treat; LESU, lesinurad; sUA, serum uric acid.

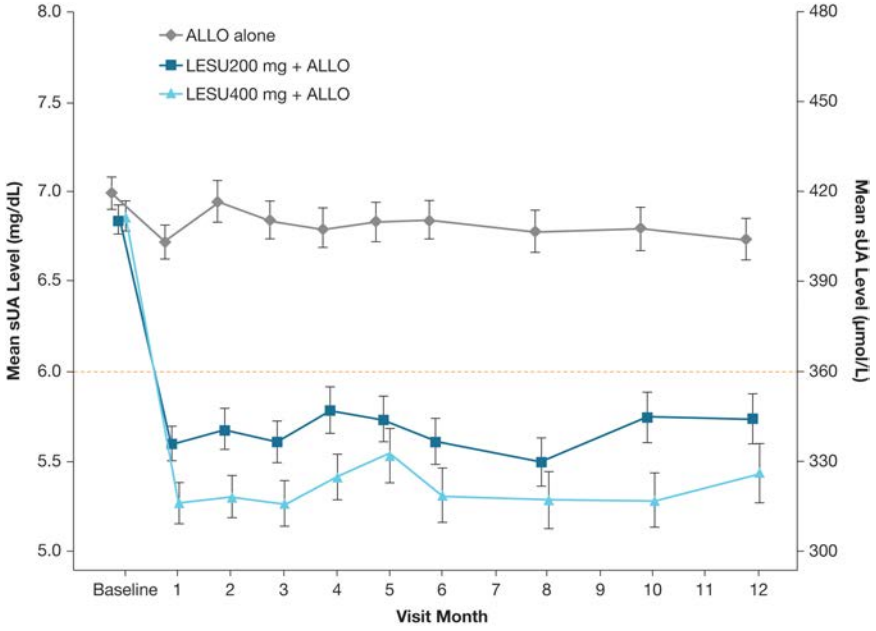


Figure 4 Graph showing the mean (SE) sUA levels by visit (observed cases, intent-to-treat population). Mean change from baseline for each active treatment group was compared with the ALLO-alone group using analysis of covariance, with p<0.001 at each time point. ALLO, allopurinol; LESU, lesinurad; sUA, serum uric acid.

the lesinurad 400 mg+allopurinol group. Non-MACE CV end points were reported in five patients (five events), two patients (two events) and no patients, respectively.

Other clinical laboratory tests and vital signs

Clinical laboratory results (excluding renal laboratory results, reported above) and urinalysis were comparable between treatment groups. Elevations in creatine kinase >5× upper limit of normal were in 5.3%, 2.0% and 3.0% of allopurinol-alone, lesinurad 200 mg+allopurinol and lesinurad 400 mg+allopurinol groups, respectively. There were no notable changes in vital signs.

DISCUSSION

Allopurinol at the 300 mg dose is frequently unable to achieve target sUA levels.[5–8] Guidelines recommend increasing the allopurinol dose above 300 mg/day to attain target sUA, but this happens rarely in practice, in part due to physician's concerns over safety of doses >300 mg.[4 6 25–29] Other management options include switching from allopurinol to febuxostat, or adding a uricosuric to allopurinol, based on evidence from earlier, small trials.[30–32] CLEAR 2 and the similarly designed CLEAR 1[15] were the initial large studies to validate a combination approach using a URAT1 inhibitor that inhibits uric acid reabsorption (ie, lesinurad) with allopurinol.

In CLEAR 2, lesinurad at both doses (200 or 400 mg) combined with continued allopurinol significantly increased the proportions of patients achieving sUA target of <6.0 mg/dL (<357 µmol/L) by month 6 (p<0.0001), with more than twice as many patients reaching goal versus allopurinol-alone therapy. Onset of sUA reduction in the lesinurad groups was rapid, with significant differences from allopurinol-alone group by first assessment at month 1. The significant increase in proportions of patients who achieved sUA target in both lesinurad+allopurinol groups versus allopurinol-alone group was sustained over the 12-month study. Consistent response rates were observed, irrespective of renal function or thiazide diuretic use.

There were no statistically significant differences favouring lesinurad treatment in the rates of gout flare requiring treatment or complete resolution of tophi, which occurred at low incidences at baseline and during study. In relation to these key secondary end points, treatment may be required for more than 12 months for the full effects to be observed.[33]

Lesinurad was generally well tolerated, particularly at the 200 mg dose, where the TEAE and serious TEAE profiles were comparable with the allopurinol-alone group. Higher TEAE incidences were seen in the lesinurad 400 mg+allopurinol group. Renal-related TEAEs occurred at similar incidences in the lesinurad 200 mg+allopurinol and allopurinol-alone groups, with a higher incidence in the lesinurad 400 mg+allopurinol than lesinurad 200 mg+allopurinol group; the lesinurad 400 mg+allopurinol group also showed a higher incidence of sCr elevation. The majority of sCr elevations resolved by the next assessment and in most cases without interruption in study medication. Mean renal function did not differ between the treatment groups both before and after treatment. The mechanism of sCr elevation associated with lesinurad may be via increased excretion of urinary uric acid, which has the potential to induce uric acid microcrystallisation in the renal tubules. Urine protein-to-creatinine ratio and urinalyses did not change during the study, suggesting that sCr elevation was not associated with renal parenchymal sequelae. Patients with unresolved sCr elevations showed no defining characteristics compared with those whose sCr elevation resolved.

Other therapies which inhibit URAT1 have been associated with development of kidney stones.[34 35] The lack of increase in kidney stone numbers during lesinurad therapy is potentially because of concomitant allopurinol use, which reduces uric acid production.[36 37] The rate of nephrolithiasis may also have been influenced by timing of lesinurad administration, as once-daily dosing in the morning increases urinary uric acid at a time when urine volume and urine pH are highest and the potential for uric acid precipitation is lowest.[38 39]

Prescribing information for the approved dose of lesinurad 200 mg recommends assessment of renal function prior to initiation of therapy and periodically thereafter, particularly in patients whose CrCl is 30–<45 mL/min, with discontinuation recommended if CrCl is persistently <30 mL/min (ie, severe renal impairment). Lesinurad is also contraindicated in subjects with end-stage renal disease, kidney transplant recipients or patients on dialysis.

Table 2 Overall summary of TEAEs (safety population)

Adverse event category, n (%)	ALLO alone (n=206)	Lesinurad 200 mg+ALLO (n=204)	Lesinurad 400 mg+ALLO (n=200)
Any TEAE	146 (70.9)	152 (74.5)	161 (80.5)
Any TEAE with RCTC toxicity grade 3 or 4	23 (11.2)	19 (9.3)	27 (13.5)
Any TEAE possibly related to randomised study medication	39 (18.9)	40 (19.6)	50 (25.0)
Any serious TEAE	8 (3.9)	9 (4.4)	19 (9.5)
Any fatal TEAE	0	0	2 (1.0)
Any TEAE leading to randomised study medication discontinuation	11 (5.3)	7 (3.4)	19 (9.5)
Any TEAE leading to study withdrawal	7 (3.4)	4 (2.0)	12 (6.0)
Individual serious TEAEs, n (%)			
Infections and infestations			
Pneumonia	0	2 (1.0)	0
Bronchopneumonia	0	0	1 (0.5)
Cellulitis	0	0	1 (0.5)
Empyema	0	1 (0.5)	0
Pyelonephritis chronic	0	0	1 (0.5)
Sinobronchitis	0	1 (0.5)	0
Abscess limb	1 (0.5)	0	0
Appendicitis	1 (0.5)	0	0
Diverticulitis	1 (0.5)	0	0
Neoplasms benign, malignant and unspecified			
Basal cell carcinoma	0	0	1 (0.5)
Gastric cancer	0	0	1 (0.5)*
Ovarian adenoma	0	1 (0.5)	0
Parathyroid tumour benign	0	1 (0.5)	0
Prostate cancer	0	0	1 (0.5)
Pancreatic neuroendocrine tumour	1 (0.5)	0	0
Metabolism and nutrition disorders			
Gout flare	0	0	2 (1.0)
Psychiatric disorders			
Depression	0	1 (0.5)	0
Dissociative disorder	0	0	1 (0.5)
Nervous system disorders			
Subarachnoid haemorrhage	1 (0.5)	0	0
Cardiac disorders			
Myocardial infarction	0	0	3 (1.5%)
Atrial fibrillation	0	1 (0.5)	0
Coronary artery disease	0	0	1 (0.5)
Intracardiac thrombus	0	0	1 (0.5)
Vascular disorders			
Hypertensive crisis	0	0	1 (0.5)
Respiratory, thoracic and mediastinal disorders			
Pulmonary oedema	0	0	1 (0.5)*
Gastrointestinal disorders			
Duodenal ulcer haemorrhage	1 (0.5)	0	1 (0.5)
Gastrointestinal haemorrhage	0	1 (0.5)	0
Musculoskeletal and connective tissue disorders			
Osteoarthritis	0	0	2 (1.0)
Arthralgia	0	1 (0.5)	0
Back pain	0	1 (0.5)	0
Flank pain	0	1 (0.5)	0
Intervertebral disc degeneration	0	0	1 (0.5)
Renal and urinary disorders			
Nephrolithiasis	0	0	2 (1.0)
Renal failure acute	1 (0.5)	0	1 (0.5)
Renal impairment	0	0	1 (0.5)
General disorders and administration site conditions			
Adverse drug reaction	0	1 (0.5)	0
Non-cardiac chest pain	0	1 (0.5)	0
Injury, poisoning and procedural complications			
Multiple drug overdose	0	1 (0.5)	0
Multiple injuries	0	1 (0.5)	0
Femur fracture	1 (0.5)	0	0

*Fatal serious TEAE.
ALLO, allopurinol; RCTC, Rheumatology Common Toxicity Criteria; TEAE, treatment-emergent adverse event.

CV comorbidities and risk factors were present in approximately 80% of patients, reflecting the high rates of CV disease in patients with gout.[40–42] The proportions of patients with TEAEs classified as CV events during study were low and similar in treatment groups. Incidences of MACE events—that is, serious CV events including CV deaths, non-fatal MI and non-fatal stroke—were similarly low. Three patients experienced MACE in the study, all receiving lesinurad 400 mg. Low rates of MACE events during gout treatment were also reported in the open-label Long-term Allopurinol Safety Study Evaluating Outcomes in Gout Patients (LASSO) study, which reported a rate of 0.58% over 6 months for MACE during allopurinol treatment (incidence rate 1.42/100 patient-years).[6]

Limitations of CLEAR 2 include the limited data on allopurinol doses >300 mg, the relatively low proportion of women enrolled, low number of patients with evaluable tophi and the relatively short-term follow-up period that limits the ability to adequately study flares and tophi. Rates of gout flares and tophus resolution over the longer term are being investigated in an extension study (NCT01808131).

In conclusion, lesinurad (200 and 400 mg), a novel SURI, in combination with allopurinol significantly increased the proportion of patients achieving the target sUA of <6.0 mg/dL (<357 μmol/L) by month 6 and other sUA end points compared with allopurinol-alone therapy. There were no statistically significant treatment-group differences favouring lesinurad for rate of gout flares or complete tophus resolution. The combination therapy was generally well tolerated, particularly at the 200 mg lesinurad dose approved by the US Food and Drug Administration and European Medicines Agency, except for higher incidences of predominantly reversible sCr elevation compared with allopurinol-alone therapy. There were no cases of unresolved sCr elevation ≥1.5× in the lesinurad 200 mg +allopurinol group, versus three unresolved cases in the allopurinol-alone group and seven in the lesinurad 400 mg

+allopurinol group. By using a dual mechanism approach to reduce sUA, combination therapy with lesinurad and allopurinol represents a treatment option for patients with gout inadequately controlled on allopurinol-alone therapy.

Author affiliations
[1]Rhumatologie, Lariboisière Hospital, and Université Paris Diderot Sorbonne Cité, Paris, France
[2]Division of Rheumatology, Duke University School of Medicine, Durham, North Carolina, USA
[3]Division of Rheumatology, University of Michigan, Ann Arbor, Michigan, USA
[4]Biometrics, Ardea Biosciences, Inc., San Diego, California, USA
[5]Research & Development, Ardea Biosciences, Inc., San Diego, California, USA
[6]Research & Development, AstraZeneca Pharmaceuticals, Gaithersburg, Maryland, USA
[7]Medical Affairs, Ardea Biosciences, Inc., San Diego, California, USA
[8]Service de rhumatologie, Université de Lausanne, Lausanne, Switzerland

Acknowledgements Editorial support for this manuscript was provided by Bill Wolvey of PAREXEL, which was funded by AstraZeneca.

Contributors TB, RTK, PPK and AS: Criterion 1: (1) substantial contributions to study conception and design and/or (2) substantial contributions to analysis and interpretation of data; criterion 2: drafting the article or revising it critically for important intellectual content and criterion 3: final approval of the version of the article to be published. JK, MF, NB, CS and SB: Criterion 1: (1) substantial contributions to study conception and design and/or (2) substantial contributions to acquisition of data and/or (3) substantial contributions to analysis and interpretation of data; criterion 2: drafting the article or revising it critically for important intellectual content and criterion 3: final approval of the version of the article to be published. SA: criterion 1: (1) substantial contributions to acquisition of data and/or (2) substantial contributions to analysis and interpretation of data; criterion 2: drafting the article or revising it critically for important intellectual content and criterion 3: final approval of the version of the article to be published.

Funding This clinical study was funded by Ardea Biosciences, a member of the AstraZeneca group. The study sponsor had a role in the design and conduct of the study; collection, management, analysis and interpretation of the data and review and approval of the manuscript.

Competing interests TB: grant/research support from Ipsen, Menarini and consultant for AstraZeneca, Ipsen, Menarini, Novartis, Savient, Sobi, Takeda and Cymabay. RTK: consultant for AstraZeneca, Crealta Pharmaceuticals and Takeda. PPK: research grant: AstraZeneca. JK (former employee), MF, NB, CS (former employee) and SB: full-time employees of Ardea Biosciences, a member of the AstraZeneca Group. SA: full-time employee of AstraZeneca Pharmaceuticals. AS: consultant for Novartis, AstraZeneca, Menarini.

Patient consent Obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Bardin T, Richette P. Definition of hyperuricemia and gouty conditions. *Curr Opin Rheumatol* 2014;26:186–91.
2 Khanna D, Fitzgerald JD, Khanna PP, *et al*. 2012 American College of Rheumatology guidelines for management of gout. Part 1: systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. *Arthritis Care Res (Hoboken)* 2012;64:1431–46.
3 Khanna PP, FitzGerald J. Evolution of management of gout: a comparison of recent guidelines. *Curr Opin Rheumatol* 2015;27:139–46.
4 Richette P, Doherty M, Pascual E, *et al*. 2016 updated EULAR evidence-based recommendations for the management of gout. *Ann Rheum Dis* Published Online First: 25 July 2016. doi: 10.1136/annrheumdis-2016-209707.
5 Becker MA, Schumacher HR, Jr., Wortmann RL, *et al*. Febuxostat compared with allopurinol in patients with hyperuricemia and gout. *N Engl J Med* 2005;353:2450–61.
6 Becker MA, Fitz-Patrick D, Choi HK, *et al*. An open-label, 6-month study of allopurinol safety in gout: the LASSO study. *Semin Arthritis Rheum* 2015;45:174–83.
7 Edwards NL. Febuxostat: a new treatment for hyperuricaemia in gout. *Rheumatology (Oxford)* 2009;48(Suppl 2):ii15–19.
8 Schumacher HR Jr, Becker MA, Wortmann RL, *et al*. Effects of febuxostat versus allopurinol and placebo in reducing serum urate in subjects with hyperuricemia and gout: a 28-week, phase III, randomized, double-blind, parallel-group trial. *Arthritis Rheum* 2008;59:1540–8.
9 Bobulescu IA, Moe OW. Renal transport of uric acid: evolving concepts and uncertainties. *Adv Chronic Kidney Dis* 2012;19:358–71.
10 Fleischmann R, Kerr B, Yeh LT, *et al*. Pharmacodynamic, pharmacokinetic and tolerability evaluation of concomitant administration of lesinurad and febuxostat in gout patients with hyperuricaemia. *Rheumatology (Oxford)* 2014;53:2167–74.
11 Girardet J-L, Miner JN. Urate crystal deposition disease and gout—new therapies for an old problem. *Annu Rep Med Chem* 2014;49:151–64.
12 Perez-Ruiz F, Hingorani V, Welp J. Efficacy and safety of RDEA594, a novel uricosuric agent, as combination therapy with allopurinol in gout patients: randomized, double-blind, placebo-controlled, phase 2 experience. *Ann Rheum Dis* 2010;69:609.
13 Perez-Ruiz F, Sundy JS, Miner JN, *et al*. Lesinurad in combination with allopurinol: results of a phase 2, randomised, double-blind study in patients with gout with an inadequate response to allopurinol. *Ann Rheum Dis* 2016;75:1074–80.
14 Shen Z, Yeh LT, Kerr B. RDEA594, a novel uricosuric agent, shows significant additive activity in combination with allopurinol in gout patients. *American Society for Clinical Pharmacology and Therapeutics*; March 2-5, 2011; Dallas, TX.
15 Saag KG, Fitz-Patrick D, Kopicko J, *et al*. Lesinurad combined with allopurinol: randomized, double-blind, placebo-controlled study in gout subjects with inadequate response to Standard of Care Allopurinol (A US-based Study). *Arthritis Rheumatol*. Published Online First: 26 Aug 2016. doi: 10.1002/art.39840
16 Wallace SL, Robinson H, Masi AT, *et al*. Preliminary criteria for the classification of the acute arthritis of primary gout. *Arthritis Rheum* 1977;20:895–900.
17 Hande KR, Noone RM, Stone WJ. Severe allopurinol toxicity. Description and guidelines for prevention in patients with renal insufficiency. *Am J Med* 1984;76:47–56.
18 Dalbeth N, Schauer C, MacDonald P, *et al*. Methods of tophus assessment in clinical trials of chronic gout: a systematic literature review and pictorial reference guide. *Ann Rheum Dis* 2011;70:597–604.
19 Gustafsson D, Unwin R. The pathophysiology of hyperuricaemia and its possible relationship to cardiovascular disease, morbidity and mortality. *BMC Nephrol* 2013;14:164.
20 Perez-Ruiz F, Martinez-Indart L, Carmona L, *et al*. Tophaceous gout and high level of hyperuricaemia are both associated with increased risk of mortality in patients with gout. *Ann Rheum Dis* 2014;73:177–82.
21 White WB, Faich G, Borer JS, *et al*. Cardiovascular thrombotic events in arthritis trials of the cyclooxygenase-2 inhibitor celecoxib. *Am J Cardiol* 2003;92:411–18.
22 Papp KA, Fonjallaz P, Casset-Semanaz F, *et al*. Analytical approaches to reporting long-term clinical trial data. *Curr Med Res Opin* 2008;24:2001–8.
23 Bellomo R, Ronco C, Kellum JA, *et al*. Acute renal failure—definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care* 2004;8:R204–12.
24 Mehta RL, Kellum JA, Shah SV, *et al*. Acute Kidney Injury Network: report of an initiative to improve outcomes in acute kidney injury. *Crit Care* 2007;11:R31.
25 Pandya BJ, Riedel AA, Swindle JP, *et al*. Relationship between physician specialty and allopurinol prescribing patterns: a study of patients with gout in managed care settings. *Curr Med Res Opin* 2011;27:737–44.
26 Harrold LR, Mazor KM, Negron A, *et al*. Primary care providers' knowledge, beliefs and treatment practices for gout: results of a physician questionnaire. *Rheumatology (Oxford)* 2013;52:1623–9.
27 Jordan KM, Cameron JS, Snaith M, *et al*. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of gout. *Rheumatology (Oxford)* 2007;46:1372–4.
28 Sarawate CA, Brewer KK, Yang W, *et al*. Gout medication treatment patterns and adherence to standards of care from a managed care perspective. *Mayo Clin Proc* 2006;81:925–34.
29 Wise E, Khanna PP. The impact of gout guidelines. *Curr Opin Rheumatol* 2015;27:225–30.
30 Perez-Ruiz F, Alonso-Ruiz A, Calabozo M, *et al*. Efficacy of allopurinol and benzbromarone for the control of hyperuricaemia. A pathogenic approach to the treatment of primary chronic gout. *Ann Rheum Dis* 1998;57:545–9.
31 Reinders MK, van Roon EN, Houtman PM, *et al*. Biochemical effectiveness of allopurinol and allopurinol-probenecid in previously benzbromarone-treated gout patients. *Clin Rheumatol* 2007;26:1459–65.
32 Takahashi S, Moriwaki Y, Yamamoto T, *et al*. Effects of combination treatment using anti-hyperuricaemic agents with fenofibrate and/or losartan on uric acid metabolism. *Ann Rheum Dis* 2003;62:572–5.

33 Schumacher HR Jr, Becker MA, Lloyd E, *et al*. Febuxostat in the treatment of gout: 5-yr findings of the FOCUS efficacy and safety study. *Rheumatology (Oxford)* 2009;48:188–94.
34 Bach MH, Simkin PA. Uricosuric drugs: the once and future therapy for hyperuricemia? *Curr Opin Rheumatol* 2014;26:169–75.
35 Perez-Ruiz F, Hernandez-Baldizon S, Herrero-Beites AM, *et al*. Risk factors associated with renal lithiasis during uricosuric treatment of hyperuricemia in patients with gout. *Arthritis Care Res (Hoboken)* 2010;62:1299–305.
36 Ettinger B, Tang A, Citron JT, *et al*. Randomized trial of allopurinol in the prevention of calcium oxalate calculi. *N Engl J Med* 1986;315:1386–9.
37 Goldfarb DS, MacDonald PA, Gunawardhana L, *et al*. Randomized controlled trial of febuxostat versus allopurinol or placebo in individuals with higher urinary uric acid excretion and calcium stones. *Clin J Am Soc Nephrol* 2013;8:1960–7.
38 Shen Z, Rowlings C, Kerr B, *et al*. Pharmacokinetics, pharmacodynamics, and safety of lesinurad, a selective uric acid reabsorption inhibitor, in healthy adult males. *Drug Des Devel Ther* 2015;9:3423–34.
39 Ngo TC, Assimos DG. Uric acid nephrolithiasis: recent progress and future directions. *Rev Urol* 2007;9:17–27.
40 Feig DI, Kang DH, Johnson RJ. Uric acid and cardiovascular risk. *N Engl J Med* 2008;359:1811–21.
41 Krishnan E, Pandya BJ, Chung L, *et al*. Hyperuricemia and the risk for subclinical coronary atherosclerosis--data from a prospective observational cohort study. *Arthritis Res Ther* 2011;13:R66.
42 Krishnan E. Gout and the risk for incident heart failure and systolic dysfunction. *BMJ Open* 2012;2:e000282.

EXTENDED REPORT

Development of the autoinflammatory disease damage index (ADDI)

Nienke M ter Haar,[1,2] Kim V Annink,[3] Sulaiman M Al-Mayouf,[4] Gayane Amaryan,[5] Jordi Anton,[6] Karyl S Barron,[7] Susanne M Benseler,[8] Paul A Brogan,[9] Luca Cantarini,[10] Marco Cattalini,[11] Alexis-Virgil Cochino,[12] Fabrizio De Benedetti,[13] Fatma Dedeoglu,[14] Adriana A De Jesus,[15] Ornella Della Casa Alberighi,[16] Erkan Demirkaya,[17] Pavla Dolezalova,[18] Karen L Durrant,[19] Giovanna Fabio,[20] Romina Gallizzi,[21] Raphaela Goldbach-Mansky,[15] Eric Hachulla,[22] Veronique Hentgen,[23] Troels Herlin,[24] Michaël Hofer,[25,26] Hal M Hoffman,[27] Antonella Insalaco,[28] Annette F Jansson,[29] Tilmann Kallinich,[30] Isabelle Koné-Paut,[31] Anna Kozlova,[32] Jasmin B Kuemmerle-Deschner,[33] Helen J Lachmann,[34] Ronald M Laxer,[35] Alberto Martini,[36] Susan Nielsen,[37] Irina Nikishina,[38] Amanda K Ombrello,[39] Seza Ozen,[40] Efimia Papadopoulou-Alataki,[41] Pierre Quartier,[42] Donato Rigante,[43] Ricardo Russo,[44] Anna Simon,[45] Maria Trachana,[46] Yosef Uziel,[47] Angelo Ravelli,[48] Marco Gattorno,[49] Joost Frenkel[3]

Handling editor Tore K Kvien

For numbered affiliations see end of article.

Correspondence to
Dr Nienke M ter Haar, Laboratory for Translational Immunology & Department of Paediatric Immunology, University Medical Centre, Lundlaan 6, Utrecht 3584EA, The Netherlands; n.m.terhaar-2@umcutrecht.nl

NMtH and KVA are joint first authors and MG and JF are joint last authors.,

Dr Ornella Della Casa Alberighi is deceased in the course of this project.

Received 21 June 2016
Revised 27 August 2016
Accepted 8 October 2016
Published Online First 3 November 2016

To cite: ter Haar NM, Annink KV, Al-Mayouf SM, *et al*. *Ann Rheum Dis* 2017;**76**:821–830.

ABSTRACT

Objectives Autoinflammatory diseases cause systemic inflammation that can result in damage to multiple organs. A validated instrument is essential to quantify damage in individual patients and to compare disease outcomes in clinical studies. Currently, there is no such tool. Our objective was to develop a common autoinflammatory disease damage index (ADDI) for familial Mediterranean fever, cryopyrin-associated periodic syndromes, tumour necrosis factor receptor-associated periodic fever syndrome and mevalonate kinase deficiency.

Methods We developed the ADDI by consensus building. The top 40 enrollers of patients in the Eurofever Registry and 9 experts from the Americas participated in multiple rounds of online surveys to select items and definitions. Further, 22 (parents of) patients rated damage items and suggested new items.
A consensus meeting was held to refine the items and definitions, which were then formally weighted in a scoring system derived using decision-making software, known as 1000minds.

Results More than 80% of the experts and patients completed the online surveys. The preliminary ADDI contains 18 items, categorised in the following eight organ systems: reproductive, renal/amyloidosis, developmental, serosal, neurological, ears, ocular and musculoskeletal damage. The categories renal/amyloidosis and neurological damage were assigned the highest number of points, serosal damage the lowest number of points. The involvement of (parents of) patients resulted in the inclusion of, for example, chronic musculoskeletal pain.

Conclusions An instrument to measure damage caused by autoinflammatory diseases is developed based on consensus building. Patients fulfilled a significant role in this process.

INTRODUCTION

Autoinflammatory diseases (AIDs) cover a spectrum of diseases, which lead to chronic or recurrent inflammation caused by activation of the innate immune system, typically in the absence of high-titre autoantibodies.[1] Over recent decades, a number of autoinflammatory diseases have been recognised, genetic defects identified and the pathogenic mechanisms elucidated.[2]

The four most common monogenic AIDs are cryopyrin-associated periodic syndromes (CAPS), familial Mediterranean fever (FMF), mevalonate kinase deficiency (MKD) and tumour necrosis factor receptor-associated periodic fever syndrome (TRAPS). In these hereditary AIDs, chronic and recurrent inflammation can lead to both acute disease and chronic irreversible damage.[3]

Targeted therapy for many AIDs has become available with blocking interleukin-1 β signalling and/or tumour necrosis factor signalling, and for many patients, control of active inflammation can be achieved. However, organ damage may have accrued in the prediagnostic or pretherapeutic phase of the illness, particularly for those with delayed diagnosis; and the control of disease activity may not be complete in every patient.[4] Therefore, many patients may still develop chronic damage from AID. This is especially true for patients for whom effective therapy is unaffordable or unavailable since many of these biological treatments are very expensive. To date, there is no validated means of assessing the long-term burden of AID available.

Currently, there is a patient-reported validated tool to quantify acute inflammatory activity in inherited periodic fevers (the autoinflammatory disease activity index); and there is a disease severity index for FMF, but by definition these do not assess

long-term damage such as hearing loss, blindness and renal failure.[5–8] Damage indices for other rheumatic diseases such as vasculitis, systemic lupus erythematosus, dermatomyositis and juvenile idiopathic arthritis have already been developed and validated.[9–13]

When devising new damage assessment tools, therapeutic toxicity must also be considered, for example, chronic glucocorticoid toxicity, which can lead to cataract, growth failure and other damaging side effects. Thus, a comprehensive damage outcome measurement tool for AID must capture chronic and potentially irreversible disorders of structure and function that have risen in patients as a result of their autoinflammatory disease and/or its treatment. The creation of such an index was a stated aim of the European Union ERANET-PRIOMEDCHILD RaDiCEA Project No. 40-41800-98-007.

The main intended purpose of the autoinflammatory disease damage index (ADDI) is to analyse the outcome of patient groups, for example, to capture and record damage in clinical trials. In addition, it may serve as an aid to physicians in assessing the needs of their patients, for example, when trying to secure funding for biological therapies. The proposed ADDI will be designed for use in the four more commonly encountered monogenic AIDs: FMF, CAPS, TRAPS and MKD. The ADDI will ideally be used as one of a set of measures to capture the disease burden for affected patients, in addition to validated measures of disease activity, disease severity and quality of life.

METHODS

We developed the ADDI by consensus building, with online surveys based on the Delphi method followed by a face-to-face consensus meeting. The Delphi method is a widely accepted and commonly used method to structurally reach consensus in a group of experts.[14]

Selection of experts and patients

The top 40 enrollers to the Eurofever Registry, a European research database for patients with AID,[15] were invited to participate as experts; another nine experts who had not participated in the European-based Eurofever Registry were recruited from the Americas. Members of this expert group participated in multiple online surveys and were invited for the face-to-face consensus meeting. In close collaboration with the Autoinflammatory Alliance,[16] we also invited 22 patients and parents of patients with FMF, CAPS, TRAPS or MKD to participate in an online survey, and an additional 3 patients to participate in the weighting of items, using the 1000minds decision-making software (see below, step 4). Inclusion criteria for selection were (1) English-speaking patients of 18 years and older or parents of a paediatric patient with FMF, MKD, CAPS or TRAPS; and (2) provision of fully informed signed consent to participate in this exercise, separately for both online surveys and interviews.

Step 1: search for possible damage items

First, a systematic literature search was performed to establish possible damage items for FMF, MKD, CAPS and TRAPS. Inclusion of articles to be considered was based on (1) all studies and case series describing symptoms and complications of more than three patients with FMF, MKD, CAPS and/or TRAPS; (2) published in English; and (3) case reports (with three or fewer patients) were included if they described significant new damage items. All data on the prevalence of the sequelae were extracted. We included all sequelae described in studies with patients with FMF, CAPS, TRAPS and MKD, which were likely to be caused by chronic inflammation or its treatment and which persisted after resolution of inflammatory episodes.

Second, we screened all items scored in the Eurofever Registry to identify new damage items not identified from the literature review. Third, we asked patients in the first online survey to propose relevant new damage items. We interviewed the patients who gave informed consent for the interviews to try to identify other relevant damage items: we asked them specifically which complications/symptoms they most fear, and which symptoms/complications create the greatest limitation of daily life. Finally, we asked experts in the first online survey for relevant new damage items (see step 2).

Step 2: multiple rounds of online surveys with experts

Four rounds of online surveys were performed as a preparation for the consensus meeting. Experts scored all potential damage items for inclusion in the index, as well as the definitions and grading of items. Experts also suggested new items, combinations of items and new options for definitions/grading. If ≥80% of the experts endorsed an item, it was included in the index. If an item reached <50% consensus, the item was excluded. In cases where 50–80% of the experts favoured inclusion, it was reconsidered in the next round. These thresholds were also used for the definitions and grading of the items.

Step 3: face-to-face consensus meeting

The 43 experts who completed one or more of the online surveys, as well as the director of the Autoinflammatory Alliance as a patient/parent representative, were invited to the consensus meeting. The first day the definition of damage and the inclusion/definitions of the items that did not reach consensus in the online surveys were discussed. On day 2, all items that reached consensus in the online surveys were refined. The results of the online surveys with experts and the patient/parent surveys and interviews were presented per item, followed by a maximum of three voting rounds and discussion. Items and definitions with 80% consensus or more were included in the ADDI. Items with no consensus after three voting rounds were excluded. After the consensus meeting, we sent a final online survey to all participants to ask whether they agreed with the items including the definitions as proposed at the consensus meeting.

Step 4: development of a scoring system

To assign an appropriate weight to each damage item, we used the 1000minds software in order to develop the scoring system of the ADDI.[17] 1000minds is a decision-making program that compares two items in order to grade the alternatives using the Potentially All Pairwise RanKings of all possible Alternatives (PAPRIKA) method.[18] Briefly, this method provides repeated comparisons between two items; the expert or patient chooses which of the two items constitutes the greater burden for patients. Each item receives a 'preference value' according to the PAPRIKA method; this reflects the importance of this item compared with all the other items. Hence, items with the greatest burden got the highest preference value and thus received most points in the ADDI.

All experts and the patients were asked to complete 1000minds. We compared the means of the patient survey and the expert survey. Differences between the overall mean and the expert mean, as well as maximising the amount of points per category, were discussed in a web conference with a small group of experts. These experts were from different continents and included both paediatric rheumatologists and rheumatologists for adults.

RESULTS

Identification of damage items from literature search and Eurofever Registry

In the literature searches, we found 1712 articles for CAPS, 632 for MKD, 2602 for FMF and 486 for TRAPS; after screening for title and abstract, 150 articles for CAPS, 87 for MKD, 251 for FMF and 55 for TRAPS remained. After screening for full text, we included 36 articles for CAPS, 9 for MKD, 54 for FMF and 8 for TRAPS; in total, 49 separate damage items were extracted from these articles (figure 1). Eight additional items extracted from the Eurofever Registry were arterial and venous thrombosis, arterial aneurysm, large vessel vasculopathy, pulmonary fibrosis, lymphatic dysplasia, camptodactyly and kyphoscoliosis. All these items were included in the online surveys with experts and patients. No new items were selected from the case reports.

Patient/parent online survey and interviews

Twenty-two patients/parents of patients provided informed consent to participate in the online surveys. Twenty-one patients (95%) completed the online survey and nine of them gave informed consent for an interview. For patient characteristics, see table 1. Patients/parents suggested 18 new damage items, including sexual dysfunction, chronic fatigue and chronic musculoskeletal pain (table 2). The five most important damage items according to patients were AA amyloidosis, joint damage, vision loss, neurological damage and renal failure. All these items were included in the preliminary ADDI.

Expert online surveys

Forty-nine experts were invited for the online surveys. The median number (range) of included patients in the Eurofever Registry for the 40 Eurofever experts was 49 (19-194) patients per expert.

All rounds were completed by >80% of the experts. Experts suggested 16 new damage items, including persistent haematuria, chronic fatigue and corneal opacities (table 3). Eight items reached consensus for inclusion in the online surveys. Forty-two items were excluded as <50% of the experts voted in favour of the item. Examples were lymphatic dysplasia, sexual dysfunction and glomerulonephritis. Sexual dysfunction was excluded because experts concluded that it would be difficult to prove a causal relation with the disease (ie, whether it can be seen as disease-associated damage); moreover, it might reflect disease activity rather than damage. Seven items were discussed in the consensus meeting as between 50% and 80% of the experts wanted to include the item. Also, 6 of the 15 definitions required further discussion in the consensus meeting.

Consensus meeting

On the first day, 31 of the 43 invited participants were able to attend the meeting. The participants discussed the items and definitions that did not reach consensus in the online survey. The participants excluded neuropathy, muscle weakness and mood disorders. Consensus was reached about all definitions that needed reconsideration. On the second day, 29 experts

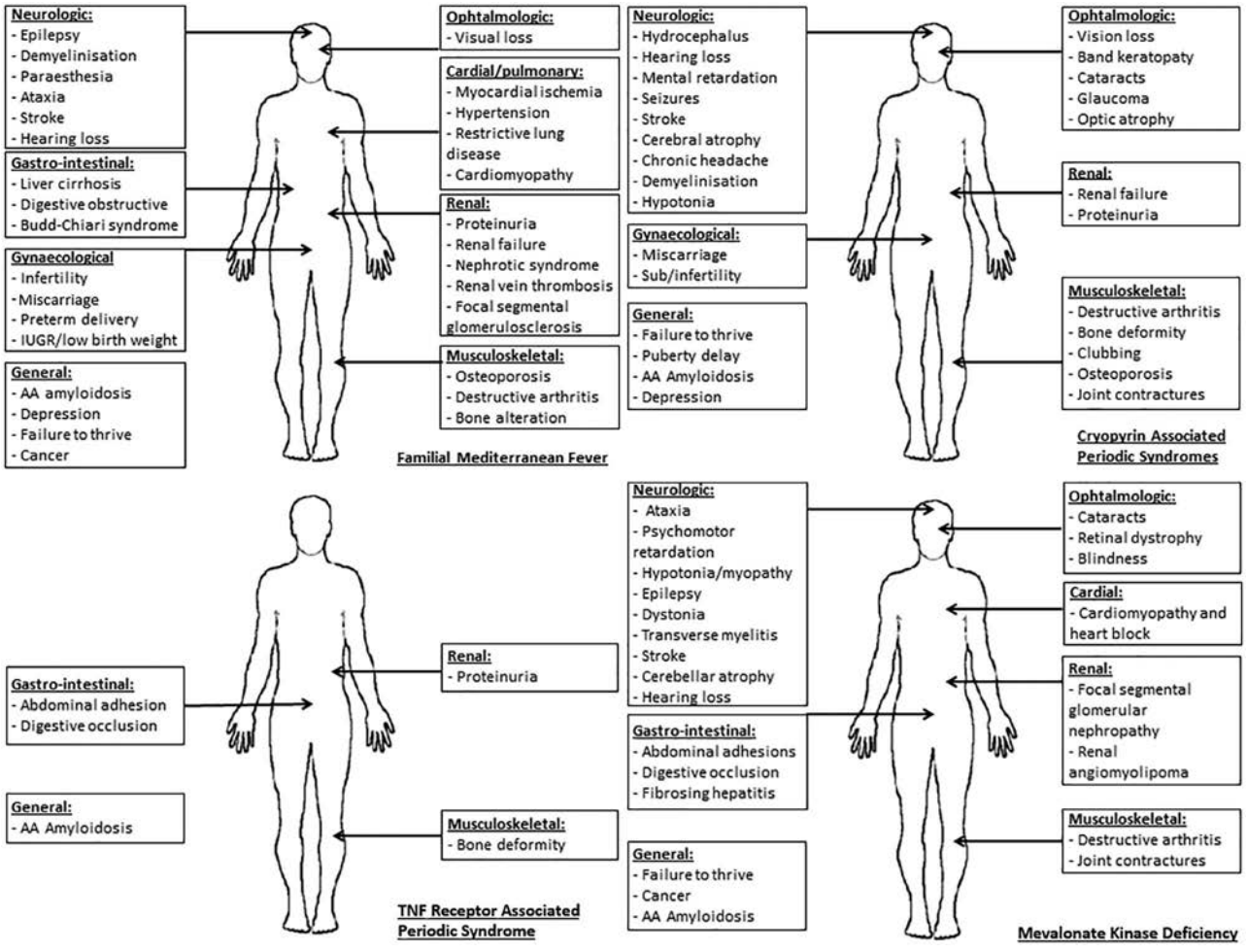


Figure 1 Damage items extracted from literature for familial Mediterranean fever,[3 19–71] cryopyrin-associated periodic syndromes,[3 72–106] tumour necrosis factor receptor-associated periodic fever syndrome[3 67 107–112] and mevalonate kinase deficiency.[3 113–119]

Table 1 Patient characteristics

	First online survey	Interviews	1000minds survey
Total no. of participants, n	21	9	14
Type of participant, n (%)			
Patients	12 (57)	3 (33)	8 (57)
Parents	9 (43)	6 (67)	6 (43)
Age, median in years (range)	28 (2–74)	15 (6–68)	29 (6–74)
Disease, n (%)			
MKD	6 (29)	1 (11)	3 (21)
TRAPS	5 (24)	3 (33)	3 (21)
CAPS	9 (43)	4 (44)	6 (43)
FMF	1 (5)	1 (11)	2 (14)
Country of residence, n (%)			
Australia	2 (10)	7 (78)	1 (7)
Canada	1 (5)	2 (22)	0 (0)
Switzerland	1 (5)	0 (0)	0 (0)
Netherlands	2 (10)	0 (0)	2 (14)
USA	15 (71)	0 (0)	10 (71)
UK	0 (0)	0 (0)	1 (7)

CAPS, cryopyrin-associated periodic syndromes; FMF, familial Mediterranean fever; MKD, mevalonate kinase deficiency; TRAPS, tumour necrosis factor receptor-associated periodic fever syndrome.

Table 2 Items suggested by patients and experts as an addition to the literature

Category	Patient suggestions	Expert suggestions
Developmental	Learning difficulties Speech developmental delay	Learning disabilities
Reproductive	Amenorrhoea Sexual dysfunction	Amenorrhoea
Neurological	Memory problems Delayed motor skill development Hand coordination problems	Hemiplegia/quadriplegia Mobility impairment
Gastrointestinal	Irritable bowel syndrome Portal hypertension	Malabsorption Portal hypertension Liver steatosis
Musculoskeletal	Craniofacial deformities	Facial deformities Muscle wasting
Ocular	Corneal haze Retinitis pigmentosa	Corneal opacity Retinitis pigmentosa
Renal		Persistent haematuria
Other	Social problems Loss of future perspective Chronic fatigue Surgeries Autonomic dysregulation Chronic pain	Weight gain Somatic growth Chronic fatigue Dysphonia

were present and refined all items that already reached consensus, including the definitions of these items. In the online survey following the consensus meeting, 35 experts agreed with almost all adaptations made in the consensus meeting. Only fatigue was finally excluded following this survey.

Most important discussions in the consensus meeting

Inclusion of infertility and amenorrhoea did not reach consensus in the online surveys, but in the consensus meeting adult rheumatologists emphasised the great burden for patients caused by infertility. After discussion, >80% of the participants agreed on including these items.

Cognitive impairment was included as an addition to developmental delay in the consensus meeting. As there is a variety of rare but severe central nervous system (CNS) complications, the participants decided to group all in one item, CNS involvement.

The group decided to replace the item abdominal adhesions with serosal scarring in order to include all potential serosal damage, for example, retroperitoneal fibrosis. Destructive arthritis and joint contractures were combined into one inclusive item, joint restriction, as movement limitation was considered the most important functional impact of both items.

Chronic headache was excluded because this item had a significant overlap with elevated intracranial pressure. Chronic musculoskeletal pain and fatigue were initially included in the consensus meeting because of the important burden for patients, albeit with a lot of discussion. Fatigue was later excluded in the final online survey because the experts agreed that although fatigue can hugely impact a patient's life, it is difficult to assess due to its subjective nature and variable relationship with disease activity.

Development of the scoring system

In total, 37 experts and 14 patients completed the 1000minds survey. The means of preference values (experts and patients) ranged from 1.5 to 7.5, in which 1.5 reflected the lowest and 7.5 the highest burden for patients. Experts and patients generally scored similar on the preference values (figure 2). A preliminary scoring system based on these preference values was presented to a panel of seven representative experts and discussed in a conference call. All items with a mean preference value of <3.5 received one point, 3.5 to 5.5 received two points (with the exception of serosal scarring, which received one point) and of >5.5 three points. Serosal scarring received one point; the experts agreed in the conference call that the consequences are less severe in comparison to other items receiving two points. Further, a maximum of points per category was defined in order to prevent double scoring of identical items. Renal/amyloidosis received a maximum amount of six points as amyloidosis often leads to renal damage. Also, the neurological and musculoskeletal categories received a decreased maximum of points because of the overlap of the items.

DISCUSSION

We developed a damage index for AID. The proposed ADDI contains 18 items. The damage items are categorised by organ system. All damage items are clearly defined and easy to score. Completing the ADDI should take approximately 5 min. The ADDI will make it possible to analyse outcomes in patient groups and compare the results of different studies, but also to systematically measure damage in a single patient.

The first key strength in the development of the ADDI is the number of worldwide experts that participated. Forty European/Middle Eastern and nine American experts were invited, with the aim of making the ADDI a global instrument. We made the selection of experts based on their clinical experience, which guarantees the capability of these experts to judge the importance of damage caused by AID. Furthermore, all online surveys were completed by >80% of the experts, which is important for both validity and acceptability of consensus statements. A high proportion of the experts attended the consensus meeting.

The second key strength is the participation of patients and parents of patients in all the steps that led to the development of the ADDI. This is important to make it a widely relevant damage index that can represent the burden for patients.

Table 3 Preliminary Autoinflammatory Disease Damage Index (ADDI) including glossary of terms

Preliminary ADDI

***Definition of damage*: Damage is defined as persistent or irreversible change in structure or function that is present for at least 6 months. Damage items should not be scored if they are attributed to ongoing disease activity. Damage may be the result of prior disease activity, complications of therapy or comorbid conditions that developed after the onset of autoinflammatory disease signs and symptoms. If damage has been present for longer than 6 months, but later resolves, it should still be scored in order to capture the damage that was present in the individual for that time period**

Damage item	Grading	Points
Reproductive		*Max. 3*
Sub/infertility		2
Amenorrhoea		1
Renal/amyloidosis		*Max. 6*
Amyloidosis	Limited amyloidosis Extensive amyloidosis	2 3
Proteinuria		1
Renal insufficiency	Moderate renal insufficiency Severe renal insufficiency	2 3
Developmental		*Max. 3*
Growth failure		2
Puberty delay		1
Serosal		*Max. 1*
Serosal scarring		1
Neurological		*Max. 6*
Developmental delay*		2
Cognitive impairment		3
Elevated intracranial pressure		2
Central nervous system involvement		3
Ears		*Max. 2*
Hearing loss	Moderate hearing loss of better ear Severe hearing loss of better ear	1 2
Ocular		*Max. 3*
Ocular involvement	Mild ocular involvement of better eye Moderate ocular involvement of better eye Severe ocular involvement of better eye	1 2 3
Musculoskeletal		*Max. 4*
Joint restriction		2
Bone deformity		2
Osteoporosis		1
Musculoskeletal pain		1

Glossary of terms

Infertility: A disease of the reproductive system defined by the failure to achieve a clinical pregnancy after ≥12 months of regular unprotected sexual intercourse, not due to known disorders in the unaffected partner.

Amenorrhoea: Primary amenorrhoea: absence of menarche at the age of 16 years or absence of menarche 5 years after thelarche in a female. Secondary amenorrhoea: absence of the menses for six consecutive months or more in a female who previously had menstrual cycles.

Limited amyloidosis: Symptomatic amyloidosis affecting one organ and confirmed by examination of tissue sections by Congo red dye or serum amyloid P component (SAP) scintigraphy.

Extensive amyloidosis: Symptomatic amyloidosis affecting more than one organ and confirmed by examination of tissue sections by Congo red dye or SAP scintigraphy.

Proteinuria: Persistent urinary protein to creatinine ratio of >20 mg/mmol in the first morning void and/or a daily protein excretion of >0.3 g/24 hours, or urine albumin to creatinine ratio of >15 mg/mmol.

Moderate renal insufficiency: Glomerular filtration rate (GFR) between 15 and 60 mL/min/1.73 m^2.

Severe renal insufficiency: GFR <15 mL/min/1.73 m^2, dialysis or transplantation.

Growth failure: Defined as the presence of at least two of the three features:
– lower than the 3rd percentile height for age
– growth velocity over 6 months lower than the 3rd percentile for age
– crossing at least two centiles (5%, 10%, 25%, 50%, 75%, 90%, 95%) on growth chart
For patients older than 18 years: Pathological short stature (eg, below 3rd percentile for normal ethnic population).

Puberty delay: A Tanner stage below –2 SDs for age.

Serosal scarring: Adhesions or fibrosis affecting pericardium, pleura, peritoneum and/or retroperitoneum, supported by imaging techniques, endoscopy or surgery.

Developmental delay: Failure to reach age-appropriate developmental milestones, including language/speech, motor, social/emotional and cognitive milestones. As soon as there is any delay in one of the development categories, this item has to be scored.*

Cognitive impairment: Requirement of special education because of cognitive impairment or IQ <70 as defined by neuropsychological assessment (eg, Wechsler Intelligence Scale for Children (WISC)) or other age-appropriate equivalents.

Elevated intracranial pressure: Signs and/or symptoms of elevated intracranial pressure supported by appropriate techniques. †

Central nervous system involvement: Focal deficits (gross and/or fine sensorimotor), diffuse deficits (eg, memory, behaviour), seizures and spinal cord symptoms.

Moderate hearing loss: Sensorineural hearing impairment confirmed by audiometry or another age-appropriate technique without requirement of hearing aids or a cochlear implant.

Severe hearing loss: Sensorineural hearing impairment confirmed by audiometry or another age-appropriate technique requiring hearing aids or a cochlear implant.

Mild ocular involvement: Ocular damage (eg, optic nerve atrophy, elevated intraocular pressure or cataract) documented by an ophthalmologist, without visual impairment.

Moderate ocular involvement: Ocular damage (eg, optic nerve atrophy, elevated intraocular pressure or cataract) documented by an ophthalmologist, resulting in visual impairment.

Severe ocular involvement: Ocular damage (eg, optic nerve atrophy, elevated intraocular pressure or cataract) documented by an ophthalmologist, resulting in legal blindness.

Joint restriction: Fixed limitation in the normal range of motion of joints, with or without destructive arthropathy or avascular necrosis.

Bone deformity: Bone deformation or overgrowth on clinical examination and/or imaging studies.

Osteoporosis: Reduced bone mineral density with vertebral collapse and/or pathological fractures confirmed with imaging, which may include bone densitometry. Requires both evidence of decreased bone density and fracture, 'low bone density' by itself is insufficient

Musculoskeletal pain: Non-inflammatory musculoskeletal pain impairing activities of daily living.

*Only for paediatric patients.
†Such as funduscopy, neuroimaging or lumbar cerebrospinal fluid (CSF) pressure measurement.

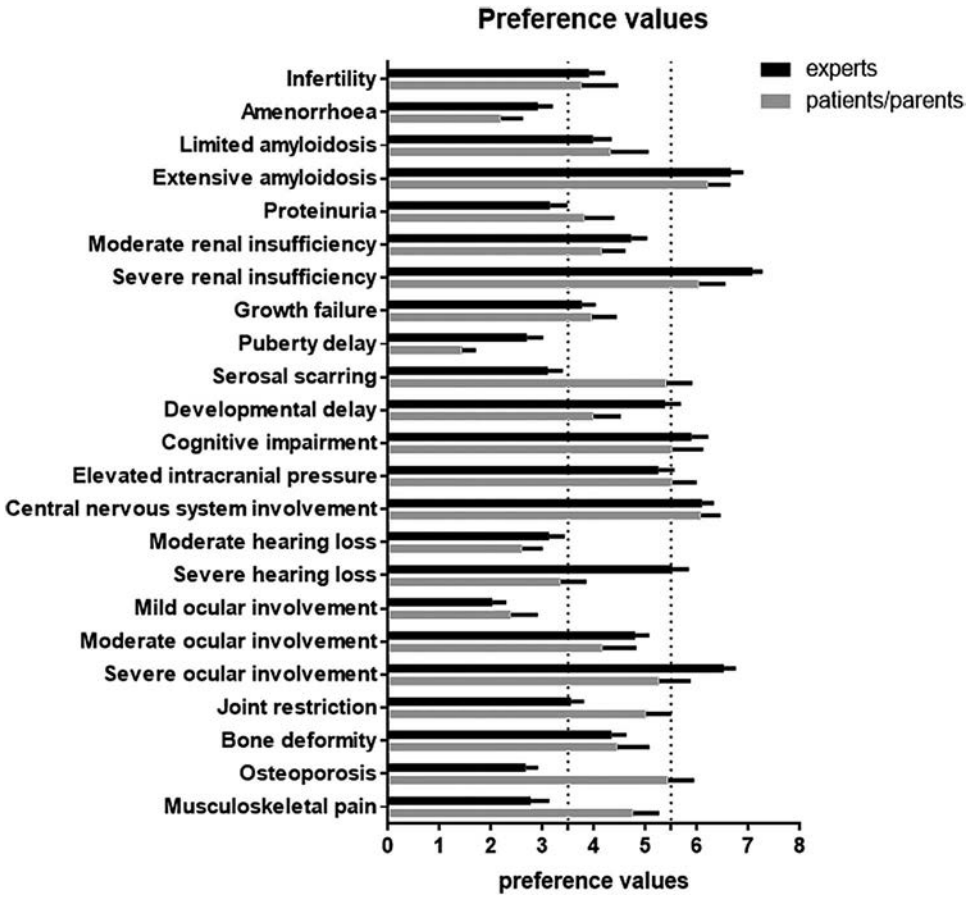


Figure 2 Scoring of the preference values from experts (black) and patients (grey), derived from the 1000minds decision-making software. A higher preference value means a higher burden for patients. The preference values range from 1.5 to 7.5, all items with a weighted mean preference value of <3.5 received one point in the Autoinflammatory Disease Damage Index (ADDI), and of >5.5 three points.

The third key strength is the methodology used to select the possible damage items. We screened for possible damage items in three ways. It was evident from the literature search that studies of long-term damage using a large sample size are extremely scarce in autoinflammatory diseases. The screening of items in the Eurofever Registry and suggestions of patients and experts were consequently valuable in developing a comprehensive set of items to asses in the online surveys.

Although many new damage items were suggested by patients and parents of patients, it might be possible that the participating patients have not suggested all possible damage items and they may not reflect the opinion of the whole patient population. Nevertheless, their contribution strengthens the process and resulted in consideration of previously neglected damage items that had not been described in the literature nor mentioned by experts, for example, chronic pain and chronic fatigue.

Patients with FMF were under-represented in this study despite attempts to recruit more patients for the 1000minds survey. Overall the amount of patients that signed informed consent as well as the response rate to surveys was lower than expected. Possible reasons might be the inclusion criterion for patients to be English speaking, the difficulty and length of the questionnaires and the informed consent procedure.

We chose to develop a general damage index limited to the four most prevalent monogenetic AIDs: FMF, CAPS, TRAPS and MKD. Based on the literature, the affected organ systems might differ in prevalence between these diseases; nevertheless, the ADDI will be a good tool to structurally score damage and covers all the important damage items for these four diseases. It would be challenging to develop the ADDI to capture damage in all AID due to the expanding number of new ultra-rare autoinflammatory diseases and their varied clinical features. An example of a recently discovered AID is the chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperature (CANDLE) syndrome. While CANDLE does share some damage items with other AID, lipodystrophy is characteristic for CANDLE,[120] but is uncommon in FMF, CAPS, TRAPS and MKD, illustrating the difficulty in developing a damage index applicable to all existing and yet to be discovered AID.

Common non-specific symptoms like chronic headache, fatigue and chronic musculoskeletal pain gave rise to intense discussions. Ultimately, only chronic musculoskeletal pain is included in the preliminary ADDI. Although patients considered these items as important in the surveys and interviews, experts thought that these items were difficult to assess objectively in daily clinical practice and found it hard to define whether these items actually reflected disease damage rather than ongoing disease activity. Nonetheless, experts acknowledged that these items have a considerable impact on the quality of life. In the future, these items might be better included in a different tool, for example, with specific items to measure quality of life.

Another difficulty in the development of the ADDI was the influence of comorbidities on the damage in AID patients. This is a common issue for all damage indices. For example, neurological impairment can be caused by the AID or by an unrelated

stroke. It is very hard to distinguish whether it is caused by independent comorbidities or the disease itself, even though we only include damage items that arose after the onset of symptoms of the AID.

In the near future, the preliminary ADDI will be validated using patient cases of FMF, CAPS, TRAPS and MKD. By this effort, we will be able to assess the validity of the ADDI in total and for the individual diseases. Furthermore, we will analyse the specificity of the ADDI items (eg, whether the damage items are not influenced by disease activity) and the grading system. Prospective validation in longitudinal cohorts will then be needed to investigate responsiveness to change over time and correlation with the burden of disease-associated damage to daily life.

In conclusion, we developed the ADDI, a universal instrument to measure persisting damage caused by chronic inflammation in the autoinflammatory diseases FMF, CAPS, TRAPS and MKD. This ADDI is based on consensus building with experts from around the world; patients and parents of patients fulfilled a significant role in this process.

Author affiliations
[1]Laboratory for Translational Immunology, University Medical Centre Utrecht, Utrecht, The Netherlands
[2]Department of Paediatric Immunology, University Medical Centre Utrecht, Utrecht, The Netherlands
[3]Department of Paediatrics, University Medical Centre Utrecht, Utrecht, The Netherlands
[4]Department of Paediatrics, King Faisal Specialist Hospital & Research Centre, Riyadh, Saudi Arabia
[5]National Paediatric Centre for Familial Mediterranean Fever and Gastroenterology Service, Arabkir Medical Centre-Institute of Child & Adolescent Health, Yerevan, Armenia
[6]Paediatric Rheumatology Unit, Hospital Sant Joan de Déu, Barcelona, Spain
[7]Division of Intramural Research and National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, USA
[8]Department of Paediatrics and Department of Rheumatology, Alberta Children's Hospital, Calgary, Canada
[9]Department of Infection, Inflammation and Rheumatology, University College London Institute of Child Health, London, UK
[10]Department of Medical Sciences, Surgery and Neurosciences, Rheumatology Unit, University of Siena, Siena, Italy
[11]Paediatric Clinic, University of Brescia and Spedali Civili di Brescia, Brescia, Italy
[12]Paediatrics Department, National Institute for Mother and Child Health Alessandrescu-Rusescu, Bucharest, Romania
[13]Division of Rheumatology, Ospedale Pediatrico Bambino Gesù, Rome, Italy
[14]Division of Immunology, Rheumatology Program, Boston Children's Hospital, Harvard Medical School, Boston, USA
[15]Translational Autoinflammatory Disease Section, NIAID, National Institutes of Health, Bethesda, USA
[16]UOSD Farmacologia Clinica e Clinical Trial—Scientific Direction, G. Gaslini Institute, Genova, Italy
[17]Department of Paediatric Rheumatology, Gulhane Military Medical Faculty, Ankara, Turkey
[18]Department of Paediatrics and Adolescent Medicine, Charles University, General University Hospital, Prague, Czech Republic
[19]Autoinflammatory Alliance, San Fransisco, USA
[20]Department of Internal Medicine, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milano, Italy
[21]Department of Paediatrics, Rheumatology, AOU G Martino, Messina, Italy
[22]Département de Médecine Interne et Immunologie Clinique, Université de Lille, Lille, France
[23]Reference centre for autoinflammatory diseases (CEREMAI), Versailles Hospital, Le Chesnay, France
[24]Department of Paediatrics, Aarhus University Hospital, Aarhus, Denmark
[25]Department of Paediatric Rheumatology, University of Lausanne, Lausanne, Switzerland
[26]Department of Paediatric Rheumatology, University Hospital of Geneva, Geneva, Switzerland
[27]Department of Paediatrics, University of California, San Diego, USA
[28]Dipartimento di Medicina Pediatrica, IRCCS Ospedale Pediatrico Bambino Gesù, Rome, Italy
[29]Department of Rheumatology&Immunology, Dr. von Hauner Childrens Hospital, Ludwig-Maximilians-University, Munich, Germany
[30]Paediatric Pneumology and Immunology and Interdisciplinary Centre for Social Paediatrics, Charité University Medicine Berlin, Berlin, Germany
[31]Department of Paediatric Rheumatology and CEREMAI, Bicêtre Hospital, APHP, University of Paris Sud, Paris, France
[32]Department of Immunology, Federal Research and Clinical Centre for Paediatric Haematology, Oncology and Immunology, Moscow, Russia
[33]Division of Paediatric Rheumatology, Department of Paediatrics, University Hospital Tuebingen, Tuebingen, Germany
[34]Division of Medicine, University College London, London, UK
[35]Department of Paediatrics and Medicine, University of Toronto and the Hospital for Sick Children, Toronto, Canada
[36]Direzione Scientifica, G Gaslini Institute, Genova, Italy
[37]Paediatric Rheumatology unit 4272, Rigshospitalet, Copenhagen, Denmark
[38]Department of Paediatric Rheumatic diseases, V.A. Nasonova Research Institute of Rheumatology, Moscow, Russia
[39]Inflammatory Disease Section, National Human Genome Research Institute, National Institutes of Health, Bethesda, USA
[40]Department of Paediatric Rheumatology, Hacettepe University, Ankara, Turkey
[41]Fourth Department of Paediatrics, Aristotle University of Thessaloniki, Thessaloniki, Greece
[42]Department of Paediatric Immunology-Hematology and Rheumatology Unit and IMAGINE Institute, Institution Necker-Enfants Malades Hospital and Paris-Descartes University, Paris, France
[43]Institute of Paediatrics, Fondazione Policlinico Universitario A. Gemelli, Università Cattolica Sacro Cuore, Rome, Italy
[44]Servicio de Inmunología y Reumatología, Hospital de Pediatría Garrahan, Buenos Aires, Argentina
[45]Internal Medicine, Radboud Expertise Centre for Immunodeficiency and Autoinflammation, Radboud University Medical Centre, Nijmegen, The Netherlands
[46]Paediatric Immunology and Rheumatology Referral Centre, first Paediatric clinic, Aristotle University of Thessaloniki, Thessaloniki, Greece
[47]Department of Paediatrics, Meir Medical Centre, Kfar Saba, Tel Aviv University, Sackler School of Medicine, Tel Aviv, Israel
[48]Institution Università degli Studi di Genova and G. Gaslini Institute, Genova, Italy
[49]UOC Pediatria 2, G. Gaslini Institute, Genova, Italy

Twitter Follow Ricardo Russo at @el_reumatologo

Acknowledgements The paper is dedicated to the memory of Dr Ornella Della Casa Alberighi, the principal investigator of RaDICEA project. The authors thank the Autoinflammatory Alliance for the collaboration and all (parents of) patients for their participation in the online surveys and interviews. Further, they would like to acknowledge Paul Hansen and Franz Ombler for providing the 1000minds decision-making software. They also thank Dr Nicolino Ruperto and the PRINTO's staff for their precious collaboration.

Contributors NMtH and KVA are joint first authors. MG and JF are joint last authors. NMtH, KVA and JF designed the study and wrote the manuscript. ODCA was the principal investigator of RADICEA. KLD contacted patients for patient recruitment. The consensus meeting was prepared with and led by AR. KLD, JF and all other authors contributed to the online surveys and/or the consensus meeting, and attributed to and approved the manuscript.

Funding The project was supported by ERANET-PRIOMEDCHILD RaDiCEA Project No. 40-41800-98-007. The Eurofever Registry was funded by the Executive Agency for Health and Consumers (EAHC, Project No. 2007332). The work was supported by an unrestricted grant by Novartis Pharma AG.

Competing interests Novartis Pharma AG financially supported the final consensus meeting. They did not have any influence on the selection of participants or on the content of the ADDI/consensus meeting or the reporting of the findings. FdB: Novartis, Novimmune, Hoffmann-La Roche, SOBI, AbbVie. LC: speaker's fee for Novartis and SOBI. MC: consultancy fees for Novartis, SOBI and Abbvie. KLD: consultancy work for SOBI and Novartis, donations, honorariums and unrestricted grants have been received by the Autoinflammatory Alliance from SOBI, Novartis, and Regeneron. RG: consultant for Abbvie. RGM: study support from SOBI, Novartis, Regeneron. VH: honorariums and educational grants from Novartis, honorariums from SOBI. MH: consultant for Novartis. HMH: consultant for Novartis and SOBI, and speaker for Novartis. TK: research grant by Novartis, speaker's bureau by Roche, BMS, Novartis and SOBI. JKD: consultant/speaker for Novartis and SOBI and has received grant support from SOBI and Novartis. RML: ad board and consultant for Abbvie and Novartis. PQ: investigator, consultant and speaker's bureau for Novartis and SOBI. MG: consultant for and unrestricted grants to Eurofever and speaker's fee from SOBI and Novartis. YU: Y. Uziel Grant/Research Support from Novartis, Consultant for Novartis, Speaker Bureau of Abbvie, Neopharm, Novartis, Roche. JF: consultant for Novartis.

Ethics approval The Medical Ethical Committee of the University Medical Centre Utrecht.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Masters SL, Simon A, Aksentijevich I, *et al*. Horror autoinflammaticus: the molecular pathophysiology of autoinflammatory disease. *Annu Rev Immunol* 2009;27:621–68.
2 Drenth JPH, van der Meer JW. Hereditary periodic fever. *N Engl J Med* 2001;345:1748–57.
3 Federici S, Sormani MP, Ozen S, *et al*. Evidence-based provisional clinical classification criteria for autoinflammatory periodic fevers. *Ann Rheumatic Dis* 2015;74:799–805.
4 Ter Haar N, Lachmann H, Ozen S, *et al*. Treatment of autoinflammatory diseases: results from the Eurofever Registry and a literature review. *Ann Rheum Dis* 2013;72:678–85.
5 Mor A, Shinar Y, Zaks N, *et al*. Evaluation of disease severity in familial Mediterranean fever. *Semin Arthritis Rheumatol* 2005;35:57–64.
6 Ozen S, Aktay N, Lainka E, *et al*. Disease severity in children and adolescents with familial Mediterranean fever: a comparative study to explore environmental effects on a monogenic disease. *Ann Rheum Dis* 2009;68:246–8.
7 Piram M, Frenkel J, Gattorno M, *et al*. A preliminary score for the assessment of disease activity in hereditary recurrent fevers: results from the AIDAI (Auto-Inflammatory Diseases Activity Index) consensus conference. *Ann Rheum Dis* 2011;70:309–14.
8 Demirkaya E, Acikel C, Hashkes P, *et al*. Development and initial validation of international severity scoring system for familial Mediterranean fever (ISSF). *Ann Rheum Dis* 2016;75:1051–6.
9 Gladman D, Ginzler E, Goldsmith C, *et al*. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheumatol* 1996;39:363–9.
10 Gladman DD, Urowitz MB, Goldsmith CH, *et al*. The reliability of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index in patients with systemic lupus erythematosus. *Arthritis Rheumatol* 1997;40:809–13.
11 Exley AR, Bacon PA, Luqmani RA, *et al*. Development and initial validation of the Vasculitis Damage Index for the standardized clinical assessment of damage in the systemic vasculitides. *Arthritis Rheumatol* 1997;40:371–80.
12 Isenberg DA, Allen E, Farewell V, *et al*. International consensus outcome measures for patients with idiopathic inflammatory myopathies. Development and initial validation of myositis activity and damage indices in patients with adult onset disease. *Rheumatology (Oxford)* 2004;43:49–54.
13 Viola S, Felici E, Magni-Manzoni S, *et al*. Development and validation of a clinical index for assessment of long-term damage in juvenile idiopathic arthritis. *Arthritis Rheumatol* 2005;52:2092–102.
14 Ruperto N, Meiorin S, Iusan SM, *et al*. Consensus procedures and their role in pediatric rheumatology. *Curr Rheumatol Rep* 2008;10:142–6.
15 Ozen S, Frenkel J, Ruperto N, *et al*. The Eurofever Project: towards better care for autoinflammatory diseases. *Eur J Pediatr* 2011;170:445–52.
16 Autoinflammatory Alliance. Autoinflammatory Disease Information-Resources-Support. http://www.autoinflammatory.org (accessed Jun 2016).
17 Hansen P, Omber F. 1000minds: decision-making software. http://www.1000minds.com (accessed Jun 2016).
18 Hansen P, Omber F. A new method for scoring multi-attribute value models using pairwise rankings of alternatives. *J Multi Criteria Decis Analysis* 2008;15:87–107.
19 Akar S, Yuksel F, Tunca M, *et al*. Familial Mediterranean fever: risk factors, causes of death, and prognosis in the colchicine era. *Medicine (Baltimore)* 2012;91:131–6.
20 Al-Wahadneh AM, Dahabreh MM. Familial Mediterranean fever in children: a single centre experience in Jordan. *Eastern Mediterr Health J* 2006;12:818–23.
21 Barakat MH, Karnik AM, Majeed HW, *et al*. Familial Mediterranean fever (recurrent hereditary polyserositis) in Arabs--a study of 175 patients and review of the literature. *Q J Med* 1986;60:837–47.
22 Bashardoust B, Maleki N. Assessment of renal involvement in patients with familial Mediterranean fever: a clinical study from Ardabil, Iran. *Intern Med J* 2014;44:1128–33.
23 Ben-Chetrit E, Ben-Chetrit A, Berkun Y. Pregnancy outcomes in women with familial Mediterranean fever receiving colchicine: is amniocentesis justified? *Arthritis Care Res* 2010;62:143–8.
24 Berkdemir Siverekli N, Sahin O, Senel S, *et al*. Bone mineral density in familial Mediterranean fever. *Rheumatology Int* 2012;32:2453–7.
25 Brik R, Gershoni-Baruch R, Shinawi M, *et al*. Pulmonary manifestations and function tests in children genetically diagnosed with FMF. *Pediatric pulm* 2003;35:452–5.
26 Brik R, Shinawi M, Kasinetz L, *et al*. The musculoskeletal manifestations of Familial Mediterranean fever in children genetically diagnosed with the disease. *Arthritis Rheumatol* 2001;44:1416–19.
27 Cefle A, Kamali S, Sayarlioglu M, *et al*. A comparison of clinical findings of familial Mediterranean fever patients with and without amyloidosis. *Rheumatology Int* 2005;25:442–6.
28 Ciftci AO, Tanyel FC, Büyükpamukçu N, *et al*. Adhesive small bowel obstruction caused by familial Mediterranean fever: the incidence and outcome. *J Pediatr Surg* 1995;30:577–9.
29 Deger SM, Ozturk MA, Demirag MD, *et al*. Health-related quality of life and its associations with mood condition in familial Mediterranean fever patients. *Rheumatol Int* 2011;31:623–8.
30 Ebrahimi-Fakhari D, Schönland SO, Hegenbart U, *et al*. Familial Mediterranean fever in Germany: clinical presentation and amyloidosis risk. *Scand J Rheumatol* 2013;42:52–8.
31 Ehrenfeld M, Brzezinski A, Levy M, *et al*. Fertility and obstetric history in patients with familial Mediterranean fever on long-term colchicine therapy. *Br J obstet Gynaecol* 1987;94:1186–91.
32 Ehrenfeld M, Levy M, Margalioth EJ, *et al*. The effects of long-term colchicine therapy on male fertility in patients with familial Mediterranean fever. *Andrologia* 1986;18:420–6.
33 Ertekin V, Selimoglu MA, Pirim I. Familial Mediterranean fever in a childhood population in eastern Turkey. *Pediatr Int* 2005;47:640–4.
34 Gedalia A, Zamir S. Neurologic manifestations in familial Mediterranean fever. *Pediatr Neurol* 1993;9:301–2.
35 Ishak GE, Khoury NJ, Birjawi GA, *et al*. Imaging findings of familial Mediterranean fever. *Clin Imaging* 2006;30:153–9.
36 Jarjour RA, Dodaki R. Arthritis patterns in familial Mediterranean fever patients and association with M694V mutation. *Mol Biol Rep* 2011;38:2033–6.
37 Kalyoncu U, Eker A, Oguz KK, *et al*. Familial Mediterranean fever and central nervous system involvement: a case series. *Medicine (Baltimore)* 2010;89:75–84.
38 Koybasi S, Atasoy HI, Bicer YO, *et al*. Cochlear involvement in Familial Mediterranean Fever: a new feature of an old disease. *Int J Pediatr Otorhinolaryngol* 2012;76:244–7.
39 Langevitz P, Livneh A, Neumann L, *et al*. Prevalence of ischemic heart disease in patients with familial Mediterranean fever. *Isr Med Assoc J* 2001;3:9–12.
40 Lidar M, Kedem R, Berkun Y, *et al*. Familial Mediterranean fever in Ashkenazi Jews: the mild end of the clinical spectrum. *J Rheumatol* 2010;37:422–5.
41 Majeed HA, Barakat M. Familial Mediterranean fever (recurrent hereditary polyserositis) in children: analysis of 88 cases. *Eur J Pediatr* 1989;148:636–41.
42 Majeed HA, Rawashdeh M. The clinical patterns of arthritis in children with familial Mediterranean fever. *QJM* 1997;90:37–43.
43 Majeed HA, Rawashdeh M, el-Shanti H, *et al*. Familial Mediterranean fever in children: the expanded clinical profile. *QJM* 1999;92:309–18.
44 Makay B, Emiroğlu N, Unsal E. Depression and anxiety in children and adolescents with familial Mediterranean fever. *Clin Rheumatol* 2010;29:375–9.
45 Migita K, Uehara R, Nakamura Y, *et al*. Familial Mediterranean fever in Japan. *Medicine (Baltimore)* 2012;91:337–43.
46 Moradian MM, Sarkisian T, Amaryan G, *et al*. Patient management and the association of less common familial Mediterranean fever symptoms with other disorders. *Genet Med* 2014;16:258–63.
47 Nobakht H, Zamani F, Ajdarkosh H, *et al*. Adult-onset familial mediterranean Fever in northwestern Iran; clinical feature and treatment outcome. *Middle East J Dig Dis* 2011;3:50–5.
48 Odabas AR, Cetinkaya R, Selcuk Y, *et al*. Familial Mediterranean fever. *South Med J* 2002;95:1400–3.
49 Ofir D, Levy A, Wiznitzer A, *et al*. Familial Mediterranean fever during pregnancy: an independent risk factor for preterm delivery. *Eur J Obstet Gynecol Reprod Biol* 2008;141:115–8.
50 Ozel AM, Demirturk L, Yazgan Y, *et al*. Familial Mediterranean fever. A review of the disease and clinical and laboratory findings in 105 patients. *Dig Liver Dis* 2000;32:504–9.
51 Polat K, Uysal IO, Senel S, *et al*. Evaluation of hearing in patients with familial Mediterranean fever. *Eur Arch Otorhinolaryngol* 2013;270:2871–4.
52 Pras E, Livneh A, Balow JE Jr, *et al*. Clinical differences between North African and Iraqi Jews with familial Mediterranean fever. *Am J Med Genet* 1998;75:216–19.
53 Reuben A, Hirsch M, Berlyne GM. Renal vein thrombosis as the major cause of renal failure in familial Mediterranean fever. *Q J Med* 1977;46:243–58.
54 Salah S, Hegazy R, Ammar R, *et al*. MEFV gene mutations and cardiac phenotype in children with familial Mediterranean fever: a cohort study. *Pediatr Rheumatol* 2014;12:5.
55 Sarikaya S, Ozdolap S, Marasli E. Spondylitis and arthritis in familial Mediterranean fever. *Turk J Rheumatol* 2012;27:241–7.
56 Savgan-Gürol E, Kasapcopur O, Hatemi S, *et al*. Growth and IGF-1 levels of children with familial Mediterranean fever on colchicine treatment. *Clin Exp Rheumatol* 2001;19:S72–75.
57 Talaat HSED, Mohamed MF, El Rifai NM, *et al*. The expanded clinical profile and the efficacy of colchicine therapy in Egyptian children suffering from familial Mediterranean fever: a descriptive study. *Ital J Pediatr* 2012;38:66.
58 Tunca M, Akar S, Onen F, *et al*. Familial Mediterranean fever (FMF) in Turkey: results of a nationwide multicenter study. *Medicine (Baltimore)* 2005;84:1–11.
59 Turgal M, Selcuk I, Ozyuncu O. Pregnancy outcome of five patients with renal amyloidosis regarding familial Mediterranean fever. *Ren Fail* 2014;36:306–8.

60 Tweezer-Zaks N, Doron-Libner A, Weiss P, *et al*. Familial Mediterranean fever and cryptogenic cirrhosis. *Medicine (Baltimore)* 2007;86:355–62.
61 Twig G, Livneh A, Vivante A, *et al*. Cardiovascular and metabolic risk factors in inherited autoinflammation. *J Clin Endocrinol Metab* 2014;99: E2123–8.
62 Twig G, Livneh A, Vivante A, *et al*. Mortality risk factors associated with familial Mediterranean fever among a cohort of 1.25 million adolescents. *Ann Rheum Dis* 2014;73:704–9.
63 Unal F, Cakir M, Baran M, *et al*. Liver involvement in children with Familial Mediterranean fever. *Dig Liver Dis* 2012;44:689–93.
64 Ureten K, Gönülalan G, Akbal E, *et al*. Demographic, clinical and mutational characteristics of Turkish familial Mediterranean fever patients: results of a single center in Central Anatolia. *Rheumatol Int* 2010;30:911–15.
65 Uthman I, Hajj-Ali RA, Arayssi T, *et al*. Arthritis in familial Mediterranean fever. *Rheumatol Int* 2001;20:145–8.
66 Uysal IÖ, Gürbüzler L, Kaya A, *et al*. Evaluation of cochlear function using transient evoked otoacoustic emission in children with Familial Mediterranean Fever. *Int J Pediatr Otorhinolaryngol* 2012;76:379–81.
67 Vergara C, Borzutzky A, Gutierrez MA, *et al*. Clinical and genetic features of hereditary periodic fever syndromes in Hispanic patients: the Chilean experience. *Clin Rheumatol* 2012;31:829–34.
68 Yasar O, Iskender C, Kaymak O, *et al*. Retrospective evaluation of pregnancy outcomes in women with familial Mediterranean fever. *J Matern -Fetal Neonatol Med* 2014;27:733–6.
69 Yazicioglu A, Turgal M, Yucel OS, *et al*. Pregnancy outcome in women with familial mediterranean fever: a retrospective analysis of 50 cases with a 10-year experience. *Turk J of Rheumatol* 2014;29:94–8.
70 Younes M, Kahn MF, Meyer O. Hip involvement in patients with familial Mediterranean fever. A review of ten cases. *Joint Bone Spine* 2002;69:560–5.
71 Zayed A, Nabil H, State O, *et al*. Subfertility in women with familial Mediterranean fever. *J Obstet Gynaecol Res* 2012;38:1240–4.
72 Ahmadi N, Brewer CC, Zalewski C, *et al*. Cryopyrin-associated periodic syndromes: Otolaryngologic and audiologic manifestations. *Otolaryngol Head Neck Surg* 2011;145:295–302.
73 Aksentijevich I, Nowak M, Mallah M, *et al*. De novo CIAS1 mutations, cytokine activation, and evidence for genetic heterogeneity in patients with neonatal-onset multisystem inflammatory disease (NOMID): a new member of the expanding family of pyrin-associated autoinflammatory diseases. *Arthritis Rheum* 2002;46:3340–8.
74 Arostegui JI, Aldea A, Modesto C, *et al*. Clinical and genetic heterogeneity among Spanish patients with recurrent autoinflammatory syndromes associated with the CIAS1/PYPAF1/NALP3 gene. *Arthritis Rheumatol* 2004;50:4045–50.
75 Caroli F, Pontillo A, D'Osualdo A, *et al*. Clinical and genetic characterization of Italian patients affected by CINCA syndrome. Rheumatology (Oxford) 2007;46:473–8.
76 Chang Z, Spong CY, Jesus AA, *et al*. Anakinra use during pregnancy in patients with cryopyrin-associated periodic syndromes (CAPS). *Rheumatology (Oxford)* 2014;66:3227–32.
77 Dodé C, Le Dû N,, Cuisset L, *et al*. New mutations of CIAS1 that are responsible for Muckle-Wells syndrome and familial cold urticaria: a novel mutation underlies both syndromes. *Am J Hum Genet* 2002;70:1498–506.
78 Dollfus H, Häfner R, Hofmann HM, *et al*. Chronic infantile neurological cutaneous and articular/neonatal onset multisystem inflammatory disease syndrome: ocular manifestations in a recently recognized chronic inflammatory disease of childhood. *Arch Ophthalmol* 2000;118:1386–92.
79 El-Darouti MA, Marzouk SA, bdel-Halim MRE. Muckle-Wells syndrome: report of six cases with hyperpigmented sclerodermoid skin lesions. *Int J Dermatol* 2006;45:239–44.
80 Goldbach-Mansky R, Dailey NJ, Canna SW, *et al*. Neonatal-onset multisystem inflammatory disease responsive to interleukin-1(beta) inhibition. *N Engl J Med* 2006;355:581–92.
81 Haas N, Küster W, Zuberbier T, *et al*. Muckle-Wells syndrome: clinical and histological skin findings compatible with cold air urticaria in a large kindred. *Br J Dermatol* 2004;151:99–104.
82 Hill SC, Namde M, Dwyer A, *et al*. Arthropathy of neonatal onset multisystem inflammatory disease (NOMID/CINCA). *Pediatr Radiol* 2007;37:145–52.
83 Hoffman HM, Wanderer AA, Broide DH. Familial cold autoinflammatory syndrome: Phenotype and genotype of an autosomal dominant periodic fever. *J Allergy Clin Immunol* 2001;108:615–20.
84 Kitley JL, Lachmann HJ, Pinto A, *et al*. Neurologic manifestations of the cryopyrin-associated periodic syndrome. *Neurology* 2010;74:1267–70.
85 Koitschev A, Gramlich K, Hansmann S, *et al*. Progressive familial hearing loss in Muckle-Wells syndrome. *Acta Oto-Laryngol* 2012;132:756–62.
86 Koné-Paut I, Lachmann HJ, Kuemmerle-Deschner JB, *et al*. Sustained remission of symptoms and improved health-related quality of life in patients with cryopyrin-associated periodic syndrome treated with canakinumab: Results of a double-blind placebo-controlled randomized withdrawal study. *Arthritis Res Ther* 2011;13:R202.
87 Kuemmerle-Deschner JB, Dembi SS, Tyrrell PN, *et al*. Challenges in diagnosing Muckle-Wells syndrome: identifying two distinct phenotypes. *Arthritis Care Res* 2014;66:765–72.
88 Kuemmerle-Deschner JB, Hachulla E, Cartwright R, *et al*. Two-year results from an open-label, multicentre, phase III study evaluating the safety and efficacy of canakinumab in patients with cryopyrin-associated periodic syndrome across different severity phenotypes. Ann Rheum Dis 2011;70:2095–102.
89 Kuemmerle-Deschner JB, Koitschev A, Ummenhofer K, *et al*. Hearing loss in Muckle-Wells syndrome. Arthritis Rheumatol 2013;65:824–31.
90 Kuemmerle-Deschner JB, Lohse P, Koetter I, *et al*. NLRP3 E311K mutation in a large family with Muckle-Wells syndrome--description of a heterogeneous phenotype and response to treatment. *Arthritis Res Ther* 2011;13:R196.
91 Kuemmerle-Deschner JB, Tyrrell PN, Koetter I, *et al*. Efficacy and safety of anakinra therapy in pediatric and adult patients with the autoinflammatory Muckle-Wells syndrome. *Arthritis Rheumatol* 2011;63:840–9.
92 Kuemmerle-Deschner JB, Tyrrell PN, Reess F, *et al*. Risk factors for severe Muckle-Wells syndrome. *Arthritis Rheumatol* 2010;62:3783–91.
93 Lainka E, Neudorf U, Lohse P, *et al*. Analysis of cryopyrin-associated periodic syndromes (CAPS) in German children: epidemiological, clinical and genetic characteristics. *Klin Padiatr* 2010;222:356–61.
94 Lepore L, Paloni G, Caorsi R, *et al*. Follow-up and quality of life of patients with cryopyrin-associated periodic syndromes treated with Anakinra. *J Pediatr* 2010;157:310–15.
95 Lequerré T, Vittecoq O, Saugier-Veber P, *et al*. A cryopyrin-associated periodic syndrome with joint destruction. *Rheumatology (Oxford)* 2007;46:709–14.
96 Leslie KS, Lachmann HJ, Bruning E, *et al*. Phenotype, genotype, and sustained response to anakinra in 22 patients with autoinflammatory disease associated with CIAS-1/NALP3 mutations. *Arch Dermatol* 2006;142:1591–7.
97 Levy R, Gérard L, Kuemmerle-Deschner J, *et al*. Phenotypic and genotypic characteristics of cryopyrin-associated periodic syndrome: a series of 136 patients from the Eurofever Registry. *Ann Rheum Dis* 2015;74:2043–9.
98 Neven B, Callebaut I, Prieur AM, *et al*. Molecular basis of the spectral expression of CIAS1 mutations associated with phagocytic cell-mediated autoinflammatory disorders CINCA/NOMID, MWS, and FCU. *Blood* 2004;103:2809–15.
99 Neven B, Marvillet I, Terrada C, *et al*. Long-term efficacy of the interleukin-1 receptor antagonist anakinra in ten patients with neonatal-onset multisystem inflammatory disease/chronic infantile neurologic, cutaneous, articular syndrome. *Arthritis Rheumatol* 2010;62:258–67.
100 Pereira AF, Pereira LB, Vale EC, *et al*. Four cases of Muckle-Wells syndrome within the same family. *An Bras Dermatol* 2010;85:907–11.
101 Prieur AM, Griscelli C, Lampert F, *et al*. A chronic, infantile, neurological, cutaneous and articular (CINCA) syndrome. A specific entity analysed in 30 patients. *Scand J Rheumatol Suppl* 1987;66:57–68.
102 Russo RA, Melo-gomes S, Lachmann HJ, *et al*. Efficacy and safety of canakinumab therapy in paediatric patients with cryopyrin-associated periodic syndrome: a single-centre, real-world experience. *Rheumatology (Oxford)* 2014;53:665–70.
103 Sibley CH, Chioato A, Felix S, *et al*. A 24-month open-label study of canakinumab in neonatal-onset multisystem inflammatory disease. *Ann Rheum Dis*2015;74:1714–19. http://dx.doi.org 10.1136/annrheumdis-2013-204877
104 Sibley CH, Plass N, Snow J, *et al*. Sustained response and prevention of damage progression in patients with neonatal-onset multisystem inflammatory disease treated with anakinra: A cohort study to determine three- and five-year outcomes. *Arthritis Rheumatol* 2012;64:2375–86.
105 Tanaka N, Izawa K, Saito MK, *et al*. High incidence of NLRP3 somatic mosaicism in patients with chronic infantile neurologic, cutaneous, articular syndrome: results of an International Multicenter Collaborative Study. *Arthritis Rheumatol* 2011;63:3625–32.
106 Tran TA, Kone-Paut I, Marie I, *et al*. Muckle-wells syndrome and male hypofertility: a case series. *Semin Arthritis Rheum* 2012;42:327–31.
107 Hull KM, Drewe E, Aksentijevich I, *et al*. The TNF receptor-associated periodic syndrome (TRAPS): Emerging concepts of an autoinflammatory disorder. *Medicine* 2002;81:349–68.
108 Lachmann HJ, Papa R, Gerhold K, *et al*. The phenotype of TNF receptor-associated autoinflammatory syndrome (TRAPS) at presentation: a series of 158 cases from the Eurofever/EUROTRAPS international registry. *Ann Rheum Dis* 2014;73:2160–7.
109 Pelagatti MA, Meini A, Caorsi R, *et al*. Long-term clinical Profile of children with the low-penetrance R92Q mutation of the tNFRSF1A gene. *Arthritis Rheumatol* 2011;63:1141–50.
110 Ravet N, Rouaghe S, Dode C, *et al*. Clinical significance of P46L and R92Q substitutions in the tumour necrosis factor superfamily 1A gene. *Ann Rheum Dis* 2006;65:1158–62.
111 Stojanov S, Dejaco C, Lohse P, *et al*. Clinical and functional characterisation of a novel TNFRSF1A c.605T>A/V173D cleavage site mutation associated with tumour necrosis factor receptor-associated periodic fever syndrome (TRAPS), cardiovascular complications and excellent response to etanercept treatment. *Ann Rheum Dis* 2008;67:1292–8.
112 Toro JR, Aksentijevich I, Hull K, *et al*. Tumor necrosis factor receptor-associated periodic syndrome: a novel syndrome with cutaneous manifestations. *Arch Dermatol* 2000;136:1487–94.

113 Drenth JP, Haagsma CJ, van der Meer JW. Hyperimmunoglobulinemia D and periodic fever syndrome. The clinical spectrum in a series of 50 patients. International Hyper-IgD Study Group. *Medicine* 1994;73:133–44.
114 Frenkel J, Houten SM, Waterham HR, *et al*. Clinical and molecular variability in childhood periodic fever with hyperimmunoglobulinaemia D. *Rheumatology (Oxford)* 2001;40:579–84.
115 Hoffmann GF, Charpentier C, Mayatepek E, *et al*. Clinical and biochemical phenotype in 11 patients with mevalonic aciduria. *Pediatrics* 1993;91:915–21.
116 Houten SM, Kuis W, Duran M, *et al*. Mutations in MVK, encoding mevalonate kinase, cause hyperimmunoglobulinaemia D and periodic fever syndrome. *Nat genet* 1999;22:175–7.
117 Poll The BT, Frenkel J, Houten SM, *et al*. Mevalonic aciduria in 12 unrelated patients with hyperimmunoglobulinaemia D and periodic fever syndrome. *J Inherit Metab Dis* 2000;23:363–6.
118 Simon A, Kremer HP, Wevers RA, *et al*. Mevalonate kinase deficiency: evidence for a phenotypic continuum. *Neurology* 2004;62:994–7.
119 Van Der Hilst JCH, Bodar EJ, Barron KS, *et al*. Long-term follow-up, clinical features, and quality of life in a series of 103 patients with hyperimmunoglobulinemia D syndrome. *Medicine (Baltimore)* 2008;87:301–10.
120 Torrelo A, Patel S, Colmenero I, *et al*. Chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperature (CANDLE) syndrome. *J Am Acad Dermatol* 2010;62:489–95.

OPEN ACCESS

EXTENDED REPORT

A randomised phase II study evaluating the efficacy and safety of subcutaneously administered ustekinumab and guselkumab in patients with active rheumatoid arthritis despite treatment with methotrexate

Josef S Smolen,[1] Sandeep K Agarwal,[2] Elena Ilivanova,[3] Xie Lillian Xu,[4] Ye Miao,[5] Yanli Zhuang,[5] Ivo Nnane,[5] Waldemar Radziszewski,[4] Andrew Greenspan,[4] Anna Beutler,[5] Daniel Baker[5]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-209831).

[1]Medical University of Vienna and Hietzing Hospital, Vienna, Austria
[2]Baylor College of Medicine, Houston, Texas, USA
[3]Leningrad Region Clinical Hospital, Saint Petersburg, Russia
[4]Janssen Research & Development, LLC, La Jolla, California, USA
[5]Janssen Research & Development, LLC, Spring House, Pennsylvania, USA

Correspondence to
Professor Josef S Smolen, Division of Rheumatology, Department of Medicine 3, Medical University of Vienna, Waehringer Guertel 18-20, Vienna A-1090, Austria; josef.smolen@wienkav.at

Received 3 May 2016
Revised 24 August 2016
Accepted 9 October 2016
Published Online First 13 January 2017

ABSTRACT

Objective Interleukin (IL)-12 and IL-23 have been implicated in the pathogenesis of rheumatoid arthritis (RA). The safety and efficacy of ustekinumab, a human monoclonal anti-IL-12/23 p40 antibody, and guselkumab, a human monoclonal anti-IL-23 antibody, were evaluated in adults with active RA despite methotrexate (MTX) therapy.

Methods Patients were randomly assigned (1:1:1:1:1) to receive placebo at weeks 0, 4 and every 8 weeks (n=55), ustekinumab 90 mg at weeks 0, 4 and every 8 weeks (n=55), ustekinumab 90 mg at weeks 0, 4 and every 12 weeks (n=55), guselkumab 50 mg at weeks 0, 4 and every 8 weeks (n=55), or guselkumab 200 mg at weeks 0, 4 and every 8 weeks (n=54) through week 28; all patients continued a stable dose of MTX (10–25 mg/week). The primary end point was the proportion of patients with at least a 20% improvement in the American College of Rheumatology criteria (ACR 20) at week 28. Safety was monitored through week 48.

Results At week 28, there were no statistically significant differences in the proportions of patients achieving an ACR 20 response between the combined ustekinumab group (53.6%) or the combined guselkumab group (41.3%) compared with placebo (40.0%) (p=0.101 and p=0.877, respectively). Through week 48, the proportions of patients with at least one adverse event (AE) were comparable among the treatment groups. Infections were the most common type of AE.

Conclusions Treatment with ustekinumab or guselkumab did not significantly reduce the signs and symptoms of RA. No new safety findings were observed with either treatment.

Trial registration number NCT01645280.

CrossMark

To cite: Smolen JS, Agarwal SK, Ilivanova E, *et al*. *Ann Rheum Dis* 2017;**76**:831–839.

INTRODUCTION

For patients with moderate-to-severe rheumatoid arthritis (RA), treatment with biologic tumour necrosis factor (TNF) inhibitors and other targeted therapies with different modes of action is often effective in reducing joint symptoms and inhibiting progression of damage.[1–7] However, many patients with RA do not respond to or lose response over time to the currently available treatments,[8 9] thus, there remains a need for novel therapies.

Interleukin (IL)-12 upregulates T helper type 1 (Th1) cell differentiation, and is the main stimulator of interferon (IFN)-γ. Increased levels of IL-12 have been detected in serum and synovial fluid of patients with RA and correlate with disease activity.[10] Also, IFN-γ is upregulated in RA synovial membranes,[11] and an IFN-γ signature can be found in peripheral mononuclear cells of patients with RA.[12] IL-23 is a member of the IL-12 cytokine family and activates Th17 cells leading to increased production of several other cytokines, including IL-17 and TNF. Various in vivo models have demonstrated that the IL-23-Th17 pathway may be involved in the development of autoimmunity, including RA.[13–15] Elevated levels of IL-23 have been observed in serum and synovial fluid of patients with RA, and serum levels decreased following anti-TNF therapy.[16–19] Furthermore, Th17 cell numbers as well as synovial IL-17 expression were found to be increased in RA.[20 21] Importantly, the extent to which RA is governed by Th1 or Th17 cells remains unclear. Therapies targeting the Th1 pathway have not been evaluated in RA, and therapies targeting IL-17 have shown variable efficacy in RA.[22]

Ustekinumab is a human monoclonal antibody targeting the IL-12/23 p40 subunit, therefore inhibiting both IL-12 and IL-23 activities, and guselkumab is an investigational monoclonal antibody targeting IL-23 specifically. Ustekinumab is effective in treating moderate-to-severe psoriasis[23 24] and active psoriatic arthritis (PsA), including inhibition of radiographic progression through 2 years.[25 26] Results of a phase II trial suggest that guselkumab may be effective in treating psoriasis;[27] guselkumab is currently being studied in phase III trials in psoriasis and a phase II trial in PsA. The current phase II trial was conducted to evaluate the efficacy and safety of ustekinumab and guselkumab in patients with active RA despite concomitant methotrexate (MTX).

PATIENTS AND METHODS

Patients

Adults (18–80 years) with a diagnosis of RA, according to the American College of Rheumatology (ACR) 1987 criteria,[28] for ≥6 months with persistent disease activity despite treatment with MTX were eligible. Persistent disease activity was defined as: ≥6/66 swollen joints and ≥6/68 tender joints and a serum C reactive protein (CRP) level ≥0.8 mg/dL. Eligibility criteria included a positive test for anticyclic citrullinated peptide antibodies or rheumatoid factor and prespecified tuberculosis screening. Patients who had received any approved or investigational biologic agent were not eligible.

Study oversight

This trial was conducted in accordance with the principles of the Declaration of Helsinki. Each patient gave written informed consent. This study was sponsored by Janssen Research & Development, LLC. The authors, some of whom are employees of the study sponsor, participated in designing the study and collecting and analysing the data. An independent data monitoring committee regularly reviewed unblinded interim safety data. All authors drafted or revised the manuscript with the assistance of a professional medical writer employed by Janssen Scientific Affairs, LLC.

Study design

This was a phase II, randomised, double-blind, placebo-controlled, parallel group trial. All patients had received MTX (10–25 mg per week) for ≥6 months before screening, and the dose was to be stable for ≥12 weeks prior to randomisation. Stable doses of concomitant glucocorticoids (<10 mg prednisone/day) and/or non-steroidal anti-inflammatory drugs (NSAIDs) and/or other analgesics for RA were permitted. Patients were randomly assigned (1:1:1:1:1) to receive placebo at weeks 0, 4 and every 8 weeks, ustekinumab 90 mg at weeks 0, 4 and every 8 weeks, ustekinumab 90 mg at weeks 0, 4 and every 12 weeks, guselkumab 50 mg at weeks 0, 4 and every 8 weeks, or guselkumab 200 mg at weeks 0, 4 and every 8 weeks through week 28. To maintain the blind, all randomised patients were to receive two 1 mL subcutaneous injections in two identical syringes prepared by an independent unblinded pharmacist at the sites at weeks 0, 4, 12, 16, 20 and 28 according to treatment assignment, including placebo injections as needed to maintain the blind, depending on the dose of ustekinumab or guselkumab assigned. The doses chosen for ustekinumab and guselkumab were based on the efficacy seen with the respective doses in previous trials of PsA and/or psoriasis.[23–27] At week 16, patients in the placebo group who had <10% improvement from baseline in both tender and swollen joint counts entered double-blind early escape and received ustekinumab 90 mg at weeks 16, 20 and 28; no treatment adjustments were made for patients randomised to the ustekinumab or guselkumab groups. The final safety follow-up visit was at week 48.

Efficacy

The primary end point was the proportion of patients who achieved at least a 20% improvement in the ACR criteria (ACR20)[29] at week 28. Other efficacy assessments included the 28-joint count Disease Activity Score using CRP (DAS28-CRP),[30] Clinical Disease Activity Index (CDAI) and Simplified Disease Activity Index (SDAI).[31] Physical function and health-related quality of life were assessed using the Health Assessment Questionnaire-Disability Index (HAQ-DI) and the 36-item Short-Form Health Survey (SF-36), respectively.[32 33]

Safety

Patients were monitored through week 48 for adverse events (AEs), clinical laboratory testing, vital signs and ECGs (weeks 0, 16 and 28).

Pharmacokinetics and immunogenicity assessments

Blood samples were collected for measuring serum ustekinumab and guselkumab concentrations and for evaluation of antidrug antibodies. Serum guselkumab concentrations were measured using a validated dissociation-enhanced lanthanide fluorescent immunoassay (lowest quantifiable guselkumab concentration: 0.04 µg/mL). Serum ustekinumab concentrations were measured by a validated electrochemiluminescence-based immunoassay method (lowest quantifiable ustekinumab concentration: 0.17 µg/mL). The presence of antibodies to ustekinumab or guselkumab in serum was determined using validated electrochemiluminescence immunoassays.

Statistical analysis

The study was powered to detect a difference in the proportion of patients achieving an ACR20 response at week 28 in the combined ustekinumab group and the combined guselkumab group compared with placebo. Combining the two dose groups of ustekinumab and guselkumab was prespecified to increase the power for detecting a difference between either ustekinumab or guselkumab and placebo as both dose groups were expected to be efficacious based on pharmacokinetic and efficacy data from other indications.[23–27] Based on a simulation of 5000 repetitions, a sample size of 50 patients per treatment group was predicted to provide approximately 74–89% power to detect an approximately 20–30% difference between a placebo plus MTX group and an active treatment (ustekinumab or guselkumab) plus MTX group. To control an overall type I error at 0.05 through the trial, a two-sided Cochran-Mantel-Haenszel test with stratification by screening CRP level (<1.5, ≥1.5 mg/dL) at a significance level of 0.025 was used in each step of a sequential testing process. For both ustekinumab and guselkumab, the combined group was compared with placebo first; if the difference from placebo was significant, pairwise comparisons between each dose group and placebo were performed. The primary end point analysis included all randomised patients grouped by randomised treatment. If a patient had data for at least one ACR component at week 28, missing component data were imputed with the last observation carried forward if baseline data were available; otherwise, missing components were considered to contribute to less than 20% improvement for ACR20 response. Patients were classified as non-responders if no ACR component data were available at week 28 or if they initiated prohibited medications (including glucocorticoids for RA), increased the MTX or glucocorticoid dose above baseline level, or discontinued the study agent for any reason. For patients who entered early escape, week-16 efficacy values were carried forward through week 28. A nominal significance level of 0.05 (two-sided) was applied to secondary end points and other analyses.

RESULTS

Patients

Data were collected from July 2012 to May 2014 at 59 sites in the USA (n=1), South America (n=72), Europe (n=197) and Asia (n=4). Five hundred and one patients were screened and

274 were randomised (placebo plus MTX, n=55; ustekinumab 90 mg every 8 weeks plus MTX, n=55; ustekinumab 90 mg every 12 weeks plus MTX, n=55; guselkumab 50 mg plus MTX, n=55; guselkumab 200 mg plus MTX, n=54) (figure 1). Baseline demographics and disease characteristics were similar across treatment groups (table 1 and online supplementary table S1). Through week 28, 22 patients discontinued the study agent; the most common reasons were lack of efficacy (n=10, 3.6%) and AEs (n=8, 2.9%) (figure 1).

Efficacy

The primary end point was not achieved; all significant differences in secondary end points are considered nominal. At week 28, an ACR20 response was achieved by 40.0% of patients in the placebo group, 53.6% in the combined ustekinumab group and 41.3% in the combined guselkumab group (p=0.101 and p=0.877, respectively); ACR20 responses in each dose group are shown in figure 2. Compared with placebo, no treatment benefit on ACR20 response was observed with ustekinumab or guselkumab in any subgroup defined by baseline demographics, disease characteristics or concomitant medications (MTX, oral glucocorticoids or NSAIDs) when compared with placebo (see online supplementary figures S1–S12). However, among ustekinumab-treated patients, a numerically greater proportion of patients in Europe achieved an ACR20 response over placebo compared with those in South America, where a high placebo response rate (58.3%) was observed. Per cent improvements in the majority of ACR components, most notably tender and swollen joint counts and physician's global assessment of disease activity, were numerically greater in the combined ustekinumab group compared with placebo at week 28; however, the effect on CRP was similar to that observed with placebo (table 2 and online supplementary table S2). No consistent improvements in ACR components were observed in the combined guselkumab group except for modest decreases in swollen and tender joint counts and physician's global assessment of disease activity, particularly in the 200 mg group.

Greater improvements from baseline in DAS28-CRP at weeks 12 and 28 were observed in the combined ustekinumab group but not in the combined guselkumab group compared with placebo at both weeks 12 and 28 (nominal p<0.05) (table 2 and online supplementary table S2). Mean changes from baseline in CDAI and SDAI were greater in the combined ustekinumab group compared with placebo at weeks 12 and 28 and in the combined guselkumab group compared with placebo at week 28 (nominal p<0.05) (table 2 and online supplementary table S2), with a trend of improvement in CDAI and SDAI Scores over time (figure 3). No improvements in mean HAQ-DI Scores were observed in both the combined ustekinumab and combined guselkumab groups at weeks 12 and 28 (table 2 and online supplementary table S2).

Safety

Through week 16, before early escape, and through week 48, the proportions of patients with at least one AE were generally comparable among the treatment groups (table 3). Through week 48, infections were the most common type of AE. Four patients had a serious infection (placebo: appendicitis (n=1); ustekinumab 90 mg every 8 weeks: urinary tract infection (n=1); guselkumab 200 mg: lobar pneumonia and gastroenteritis (n=1 each)), with no apparent difference among the treatment groups. There were no cases of tuberculosis or opportunistic infections. Overall, gastrointestinal AEs were more common in the ustekinumab 90 mg every-12-week group and blood and lymphatic disorder AEs were more common in the guselkumab 200 mg group compared with placebo (5 (9.1%) vs

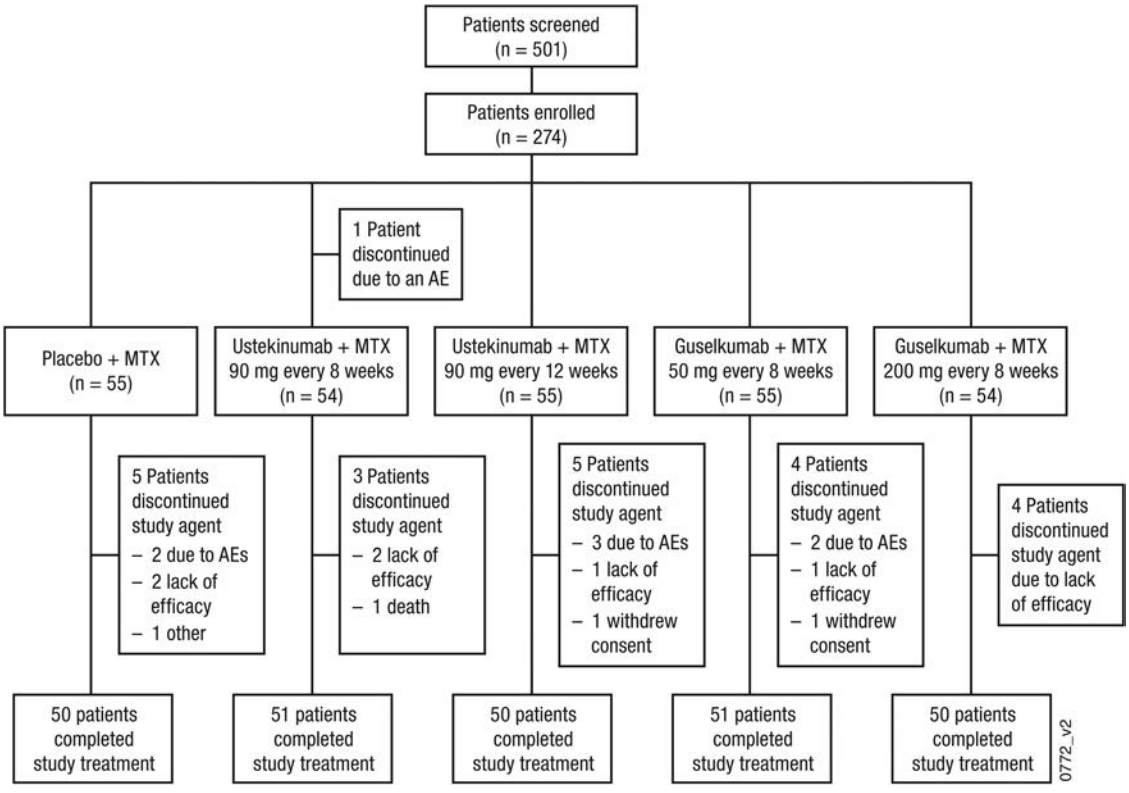


Figure 1 Patient disposition through week 28. AE, adverse event; MTX, methotrexate.

Table 1 Baseline demographic and disease characteristics

		Ustekinumab+MTX		Guselkumab+MTX		Total
	Placebo+MTX (N=55)	90 mg every 8 weeks (N=55)	90 mg every 12 weeks (N=55)	50 mg every 8 weeks (N=55)	200 mg every 8 weeks (N=54)	(N=274)
Characteristic						
Demographics						
Female sex, n (%)	48 (87.3)	46 (83.6)	47 (85.5)	45 (81.8)	42 (77.8)	228 (83.2)
Age, years	51.1±10.6	50.8±13.0	51.4±13.6	49.9±12.9	54.6±11.3	51.5±12.3
Disease duration, years	8.5±8.7	5.6±5.5	6.8±5.9	6.1±7.1	8.9±9.6	7.2±7.6
Concomitant medications						
MTX dose, mg/week	14.5±4.6	14.8±4.2	14.9±4.9	15.6±3.6	14.5±4.6	14.9±4.4
Oral glucocorticoids, n (%)	30 (54.5)	33 (60.0)	30 (54.5)	37 (67.2)	35 (64.8)	165 (60.2)
Disease characteristics						
SJC (0–66)	14.7±6.5	15.2±8.6	17.2±9.3	15.5±6.6	17.6±9.1	16.0±8.1
TJC (0–68)	26.7±11.3	26.4±14.2	27.4±12.3	26.1±12.1	28.0±13.7	26.9±12.7
Patient's assessment of pain, cm	6.4±1.9	6.6±2.0	6.5±2.2	6.6±2.1	6.5±1.9	6.5±2.0
Patient's global assessment, cm	6.5±1.8	6.8±1.9	6.8±2.0	6.8±1.7	6.7±1.7	6.7±1.8
Physician's global assessment, cm	6.8±1.3	6.3±1.3	6.4±1.5	6.6±1.6	6.7±1.4	6.5±1.4
HAQ-DI (0–3)	1.7±0.5	1.8±0.6	1.7±0.6	1.7±0.7	1.8±0.6	1.7±0.6
CRP, mg/dL (ULN ≤0.287 mg/dL)	1.9±1.6	2.3±2.5	2.0±2.2	2.3±2.3	2.3±2.2	2.2±2.2
DAS28-CRP	6.1±0.8	6.0±0.8	6.1±0.7	6.1±0.8	6.1±0.9	6.1±0.8
CDAI	41.9±11.0	40.2±10.9	43.2±11.0	41.1±10.6	42.8±13.0	41.8±11.3
SDAI	43.8±11.2	42.6±11.1	45.2±10.9	43.4±11.4	45.1±13.7	44.0±11.7
Rheumatoid factor, n (%)	48 (87.3)	47 (87.0)	51 (92.7)	53 (96.4)	50 (92.6)	249 (91.2)
Anti-CCP, n (%)	53 (96.4)	47 (87.0)	51 (92.7)	53 (96.4)	53 (98.1)	257 (94.1)

Data presented as mean±SD unless otherwise noted. No statistically significant differences (α=0.05) were observed among treatment groups.
CCP, cyclic citrullinated peptide; CDAI, Clinical Disease Activity Index; CRP, C reactive protein; DAS28-CRP, 28-joint count Disease Activity Score using CRP; HAQ-DI, Health Assessment Questionnaire-Disability Index; MTX, methotrexate; SDAI, Simplified Disease Activity Index; SJC, swollen joint count; TJC, tender joint count; ULN, upper limit of normal.

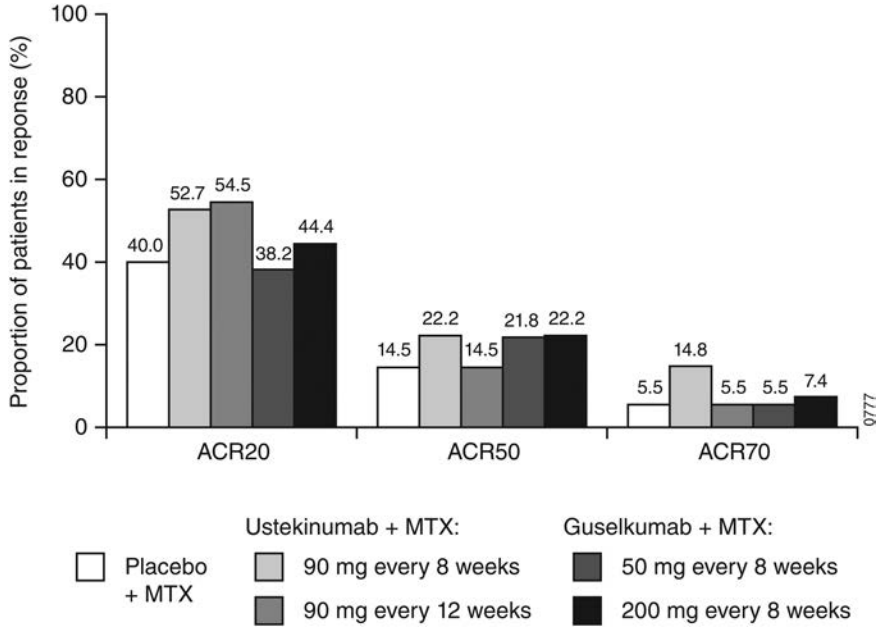


Figure 2 Proportions of patients with an ACR20, ACR50 or ACR70 response at week 28. ACR20 response includes all randomised patients. ACR50 and ACR70 responses include patients who received ≥1 dose of study agent. ACR20/50/70, ≥20%/50%/70% improvement in the American College of Rheumatology criteria; MTX, methotrexate.

2 (3.6%) and 5 (9.3%) vs 1 (1.8%), respectively). Few patients had an injection site reaction (table 3); none were serious or severe, and none led to discontinuation of the study agent.

The incidence of patients with at least one serious AE (SAE) through week 48 was low and similar across treatment groups (table 3). Two malignancies occurred, both during the follow-up period after week 28: squamous cell carcinoma of the lung (male, aged 59 years, non-smoker) in the ustekinumab 90 mg every-12-week group and breast cancer (female, aged 62 years) in the guselkumab 200 mg group. One death occurred in a woman aged 61 years in the ustekinumab 90 mg every-8-week group who experienced syncope (16 days after receiving the third administration of study drug) and was admitted to the hospital, with pulmonary embolism or pneumonia considered at admission; the exact cause of death could not be confirmed. Other SAEs were unstable angina and

Table 2 Secondary efficacy assessments at week 28

		Ustekinumab+MTX			Guselkumab+MTX		
	Placebo+MTX (N=55)	90 mg every 8 weeks (N=55)	90 mg every 12 weeks (N=55)	Combined	50 mg every 8 weeks (N=55)	200 mg every 8 weeks (N=54)	Combined
Patients, n	55	54	55	109	55	54	109
ACR20, n (%)	22 (40.0)	29 (53.7)	30 (54.5)	59 (54.1)	21 (38.2)	24 (44.4)	45 (41.3)
ACR50, n (%)	8 (14.5)	12 (22.2)	8 (14.5)	20 (18.3)	12 (21.8)	12 (22.2)	24 (22.0)
ACR70, n (%)	3 (5.5)	8 (14.8)	3 (5.5)	11 (10.1)	3 (5.5)	4 (7.4)	7 (6.4)
Per cent change in ACR core components, median (IQR)							
SJC	−26.7 (−75.0, 7.1)	−65.2 (−91.2, −28.6)	−71.9 (−86.7, −40.0)	−69.7 (−87.5, −34.8)	−50.0 (−86.7, −25.0)	−58.6 (−86.7, −40.0)	−57.1 (−86.7, −27.3)
TJC	−23.7 (−68.0, 13.6)	−43.7 (−79.2, −16.7)	−50.0 (−72.2, −25.0)	−45.8 (−75.0, −21.4)	−50.0 (−77.8, −15.8)	−45.0 (−71.4, −20.8)	−50.0 (−73.3, −20.5)
Pain, VAS	−25.8 (−56.5, 11.1)	−31.9 (−48.5, −8.8)	−20.8 (−47.7, 2.2)	−23.5 (−48.3, −5.7)	−15.6 (−45.3, 2.5)	−19.9 (−38.5, 1.2)	−19.7 (−40.0, 2.0)
Patient's global assessment of disease activity	−22.8 (−50.0, 4.6)	−31.0 (−58.3, −12.2)	−25.0 (−46.8, −4.2)	−30.1 (−49.3, −6.3)	−17.0 (−58.6, 3.7)	−16.7 (−38.9, 1.2)	−16.9 (−41.3, 2.4)
Physician's global assessment of disease activity	−26.4 (−59.4, −5.7)	−41.4 (−82.0, −16.7)	−49.4 (−64.7, −25.7)	−45.5 (−73.9, −23.3)	−52.0 (−71.9, −24.6)	−44.8 (−68.1, −20.4)	−49.2 (−71.0, −20.4)
HAQ-DI	−14.3 (−36.8, 6.3)	−21.1 (−57.1, −4.5)	−20.0 (−40.0, −7.1)	−20.0 (−45.5, −5.9)	−26.3 (−45.0. 0.0)	−18.8 (−45.0, 5.6)	−23.8 (−45.0, 0.0)
CRP	−28.4 (−60.6, 37.8)	−24.8 (−76.6, 26.1)	−29.1 (−69.1, 50.5)	−24.8 (−72.8, 40.8)	0.0 (−40.2, 141.2)	−36.0 (−69.3, 51.4)	−7.2 (−55.3, 88.4)
DAS28-CRP change from baseline, least squares mean (95% CI)	−0.9 (−1.3, −0.6)	−1.5* (−1.9, −1.2)	−1.5* (−1.9, −1.1)	−1.5** (−1.8, −1.3)	−1.4 (−1.8, −1.1)	−1.2 (−1.5, −0.9)	−1.3 (−1.6, −1.1)
DAS28-CRP response, n (%)	24 (43.6)	36 (66.7)	33 (60.0)	69 (63.3)	31 (56.4)	32 (59.3)	63 (57.8)
CDAI change from baseline, mean±SD	−11.3±16.4	−17.2±16.8	−19.9±10.9**	−18.6±14.2**	−16.7±12.8	−18.6±14.9*	−17.6±13.8*
SDAI change from baseline, mean±SD	−11.4±17.0	−17.9±17.5	−20.4±11.6**	−19.2±14.8**	−16.5±13.1	−19.1±15.5*	−17.8±14.3*
Patients in remission, n (%)	0	3 (5.6)	1 (1.8)	4 (3.7)	0	1 (1.9)	1 (0.9)
HAQ-DI change from baseline, least squares mean (95% CI)	−0.3 (−0.4, −0.1)	−0.4 (−0.6, −0.3)	−0.5 (−0.6, −0.3)	−0.5 (−0.6, −0.4)	−0.4 (−0.5, −0.2)	−0.4 (−0.6, −0.3)	−0.4 (−0.5, −0.3)

*p<0.05; **p<0.01.
ACR20/50/70, ≥20%/50%/70% improvement in the American College of Rheumatology criteria; CDAI, Clinical Disease Activity Index; CRP, C reactive protein; DAS28-CRP; 28-joint count Disease Activity Score with CRP; HAQ-DI, Health Assessment Questionnaire-Disability Index; MTX, methotrexate; SDAI, Simplified Disease Activity Index; SJC, swollen joint count; TJC, tender joint count; VAS, visual analogue scale.

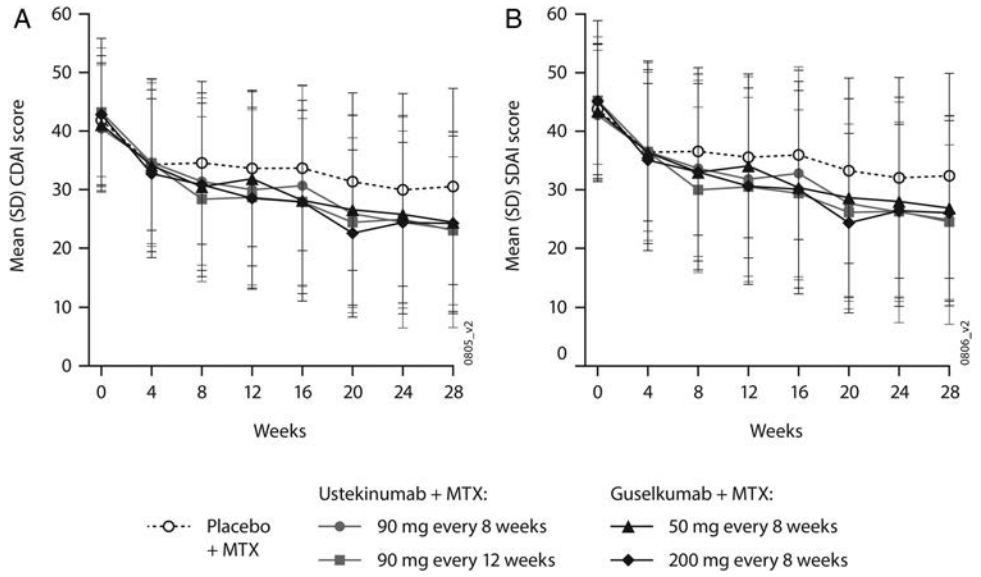


Figure 3 Mean Clinical Disease Activity Index (CDAI; panel A) and Simplified Disease Activity Index (SDAI; panel B) Scores through week 28. MTX, methotrexate.

Table 3 Adverse events summary through week 48

		Ustekinumab+MTX			Guselkumab+MTX		
	Placebo+MTX (N=55)	90 mg every 8 weeks (N=55)	90 mg every 12 weeks (N=55)	Combined	50 mg every 8 weeks (N=55)	200 mg every 8 weeks (N=54)	Combined
Through week 16							
Patients, n	55	54	55	109	55	54	109
Mean exposure, weeks	15.8	16.3	15.8	16.0	16.3	16.4	16.4
Patients with ≥1 AE, n (%)	21 (38.2)	22 (40.7)	24 (43.6)	46 (42.2)	16 (29.1)	21 (38.9)	37 (33.9)
Patients with ≥1 SAE, n (%)	1 (1.8)	2 (3.7)	2 (3.6)	4 (3.7)	0	1 (1.9)	1 (0.9)
Through week 48							
Patients, n	55	54	55	125*	55	54	109
Mean exposure, weeks	23.7	27.8	26.9	25.4	28.2	28.0	28.1
Patients with ≥1 AE, n (%)	25 (45.5)	26 (48.1)	30 (54.5)	63 (50.4)	20 (36.4)	27 (50.0)	47 (43.1)
Injection site reactions through week 28, n (%)	0	0	1 (1.8)	1 (0.8)	1 (1.8)	1 (1.9)	2 (1.8)
Infections, n (%)	16 (29.1)	13 (24.1)	21 (38.2)	37 (29.6)	12 (21.8)	13 (24.1)	25 (22.9)
Common AEs, n (%)							
Nasopharyngitis	3 (5.5)	5 (9.3)	4 (7.3)	10 (8.0)	3 (5.5)	4 (7.4)	7 (6.4)
Influenza	3 (5.5)	1 (1.9)	3 (5.5)	4 (3.2)	3 (5.5)	3 (5.6)	6 (5.5)
Worsening of RA	1 (1.8)	2 (3.7)	5 (9.1)	8 (6.4)	2 (3.6)	4 (7.4)	6 (5.5)
Headache	3 (5.5)	2 (3.7)	5 (9.1)	8 (6.4)	2 (3.6)	3 (5.6)	5 (4.6)
Hypertension	3 (5.5)	4 (7.4)	2 (3.6)	7 (5.6)	1 (1.8)	1 (1.9)	2 (1.8)
Back pain	1 (1.8)	0	0	0	3 (5.5)	1 (1.9)	4 (3.7)
Anaemia	1 (1.8)	3 (5.6)	0	3 (2.4)	1 (1.8)	3 (5.6)	4 (3.7)
Patients with ≥1 SAE, n (%)	3 (5.5)	4 (7.4)	3 (5.5)	8 (6.4)	0	3 (5.6)	3 (2.8)
Patients with ≥1 serious infection, n (%)	1 (1.8)	1 (1.9)	0	1 (0.8)	0	2 (3.7)	2 (1.8)

*Includes 16 patients who entered early escape at week 16.
AE, adverse event; MTX, methotrexate; RA, rheumatoid arthritis; SAE, serious adverse event.

worsening of RA (placebo); sciatica, anaemia, concussion and shock (ustekinumab 90 mg every 8 weeks); and ileus and transient ischaemic attack (ustekinumab 90 mg every 12 weeks). All SAEs were singular events, and no specific pattern of association between SAEs and active treatments was identified.

There were no clinically meaningful changes in vital signs or ECG findings. Most chemistry and haematology abnormalities were mild to moderate; occurrences of toxicity grades >2 were transient and not clinically significant.

Pharmacokinetics

Median trough serum levels of ustekinumab reached steady state by week 12 (1.59 µg/mL) in the every-8-week group and by week 16 (0.61 µg/mL) in the every-12-week group, and were

maintained through week 28 (1.77 μg/mL and 0.54 μg/mL, respectively). No clear correlation between trough serum ustekinumab concentrations and ACR20 response at week 28 was observed.

Median trough serum levels of guselkumab reached steady state by week 20 and were maintained through week 28 (0.18 μg/mL and 0.73 μg/mL in the 50 mg every-8-week and 200 mg every-8-week groups, respectively).

Immunogenicity

Through week 48, 7 out of 123 (5.7%) ustekinumab-treated patients with appropriate samples tested positive for antibodies to ustekinumab; four had neutralising antibodies. In both ustekinumab groups, serum ustekinumab concentrations were generally lower in patients who tested positive for antibodies to ustekinumab compared with those who tested negative.

Through week 48, 12 out of 106 (11.3%) guselkumab-treated patients with appropriate samples tested positive for antibodies to guselkumab; none had neutralising antibodies. Serum guselkumab concentrations were generally comparable between patients who tested positive for antibodies to guselkumab and those who tested negative.

DISCUSSION

The relative contributions of IL-12 and/or IL-23 pathways to the pathophysiology of RA are not well understood. This trial was undertaken to evaluate the safety and efficacy of ustekinumab (anti IL-12/23p40 antibody) and guselkumab (anti-IL-23 antibody) in patients with active RA despite MTX therapy. The primary end point (ACR20 at week 28) was not met for either ustekinumab or guselkumab. While some numerical trends towards improvement were consistently observed in several secondary efficacy measures in the ustekinumab treatment groups compared with the placebo, the reductions in composite disease activity measures DAS28-CRP, SDAI and CDAI were relatively small. No consistent evidence of efficacy in RA was observed with guselkumab in this study. No treatment effect was observed with ustekinumab or guselkumab on CRP levels in patients with RA.

The large placebo effect on ACR20 response observed at week 12 (29.1%) and week 28 (40.0%) may have made it more difficult to demonstrate efficacy for the active treatments, which may be a limitation for this study. However, it should be noted that placebo response rates have been quite high in several recent trials of similar populations, such as REALISTIC (ACR20 response: 26% at week 12[34]) or MOBILITY (ACR20 response: 46% at week 12 and 33% at week 24[35 36]); however, while the placebo rates were in the order of those observed in our trial, the response rates with the active medications were in the order of 51% to 72% and thus much higher and even up to double those observed here. The small sample size in this phase II study could also be a limitation. However, the totality of the data, including ACR 20/50/70 response over time and lack of effect on CRP, an objective measure, suggested ustekinumab and guselkumab did not have significant, clinically meaningful effects on the signs and symptoms and the inflammatory markers of RA in this patient population with moderate-to-severe disease.

As both guselkumab (which selectively inhibits IL-23) and ustekinumab (which blocks IL-12/IL-23) failed to demonstrate robust efficacy in this study, these results suggest that activation of Th17 cells may not play a major role in established RA. These data are in stark contrast to those obtained in psoriasis with both monoclonal antibodies and in PsA with ustekinumab. Indeed, the efficacy of ustekinumab in psoriasis[23 24] and PsA,[25 26] and of guselkumab in psoriasis[27] is consistent with the effects of IL-17 inhibition in both diseases. Notably, the lack of robust efficacy in patients with RA following treatment with guselkumab or ustekinumab in the present trial is also consistent with previous studies on IL-17 inhibition in patients with moderate-to-severe RA, which showed modest efficacy with secukinumab[37] and ixekizumab[38] and no efficacy with brodalumab.[39] Overall, these findings point to differences in the immunopathology of active RA and PsA joint disease; in PsA, the IL-23/IL-17-mediated Th17 pathway may play a more important role.

While the lack of demonstrable efficacy for both ustekinumab and guselkumab in patients with RA suggests that Th17 cells may only play a minor role in established RA, the effect of ustekinumab on Th1 cells suggests that Th1 cells play a limited role in the pathophysiology of RA at this (established) stage of the disease. Alternatively, Th1 cell function may be blocked insufficiently at the ustekinumab dose level evaluated in this study. However, in a previous study, evidence of inhibition of IFNγ production was shown in a subset of patients with psoriasis who had ≥75% improvement in Psoriasis Area and Severity Index after receiving a single administration of ustekinumab.[40]

Interestingly, abatacept, a T cell activation inhibitor, has shown robust efficacy on joint symptoms in both RA[41] and PsA,[42] while an effect on skin psoriasis was not observed in a phase II PsA Study.[42] However, with the exception of abatacept, no T cell directed therapy has hitherto shown efficacy in RA.[43] Also, it is not clear at present if the major mode of action of abatacept in RA is mainly related to inhibition of T cell activation or due to other mechanisms.[44–46] Irrespective of abatacept's mode of action, the difference in efficacy profiles of ustekinumab, guselkumab and abatacept as well as direct IL-17 inhibitors in RA,[37 39 47] PsA[48 49] and psoriasis[50 51] suggest that the IL-23/IL-17-mediated Th17 pathways do not play an important role in RA joint inflammation.

An alternative explanation could be that the highest ustekinumab and guselkumab exposures achieved in this study may not have been adequate for efficacy in RA. However, this is unlikely, as the serum exposures of ustekinumab and guselkumab observed in this study were generally consistent with those observed in different patient populations (eg, psoriasis and PsA) that demonstrated efficacy. Furthermore, no consistent dose response in efficacy was observed for either ustekinumab or guselkumab, and there was no clear correlation between steady state trough serum concentrations of ustekinumab or guselkumab with the proportion of patients who achieved an ACR20 response.

The safety profiles of ustekinumab and guselkumab were consistent with earlier studies in other patient populations.[23–27] Overall, the incidence of AEs in the ustekinumab and guselkumab groups was comparable with the placebo group through week 16. The incidence of SAEs was low and was similar among the treatment groups, with no specific pattern of association between SAEs and active treatments. Four patients (one receiving placebo, one receiving ustekinumab, two receiving guselkumab) had a serious infection. No opportunistic infections or cases of tuberculosis were reported. Two malignancies (one in a patient who received ustekinumab and one in a patient who received guselkumab) and one death occurred (ustekinumab group). Overall, no new safety risk or particular pattern of event clustering was evident.

In summary, patients with active RA despite prior MTX did not demonstrate any clinically meaningful improvement in the signs and symptoms of RA following treatment with ustekinumab or guselkumab despite the clear benefit of ustekinumab in

both psoriasis and PsA and robust phase II data suggesting efficacy of guselkumab in psoriasis. Our results suggest that, in contrast to psoriasis and PsA, Th17 cells, as well as Th1 cells, may not play a major role in the pathogenesis of active established RA. Additional research is needed to fully understand the relative roles of IL-12 and IL-23 in RA.

Acknowledgements The authors thank Mittie Doyle, MD, and Alan Mendelsohn, MD, formerly of Janssen Research & Development, LLC, and Yasmine Wasfi, MD, Surekha Mudivarthy, PhD, and Weichun Xu, PhD, of Janssen Research & Development, LLC, for their contributions to this study. The authors also thank Rebecca E Clemente, PhD, of Janssen Scientific Affairs, LLC, for writing support.

Investigators: *Argentina:* Horacio Oscar Venarotti, Buenos Aires; Rodolfo Ariel Pardo Hidalgo, San Juan; Guillermo Tate, Buenos Aires; Alberto Spindler, Tucumán; María Alicia Lázaro, Buenos Aires. *Bulgaria:* Zlatimir Kolarov, Sofia; Boycho Oparanov, Sofia; Anastas Batalov, Plovdiv. *Chile:* Sonia Arriagada, Osorno. *Colombia:* John Londono, Chia; William Otero, Bucaramanga; Jose Fernando Molina, Medellin; Patricia Velez, Bogota; Maria Claudia Diaz, Bogota; Diego Luis Saaibi, Bucaramanga; Beatriz Arana, Cali; Luis Fernando Pinto, Medellin. *Czech Republic:* Gabriela Simkova, Kladno; Libor Novosad, Hlucin; Eva Dokoupilova. *Hungary:* Edit Drescher, Veszprém; Bernadette Rojkovich, Budapest; Edit Agnes Toth, Gödöllő, Lilla Nafradi, Szombathely; Kata Kerekes, Szekesfehervar. *Poland:* Piotr Lesacayński, Poznan; Artur Racewicz, Bialystok; Anna Zubrzycka-Sienkiewicz, Warszawa; Przemyslaw Kotyla, Sosnowiec; Mariusz Korkosz, Krakow; Malgorzata Sochocka-Bykowska, Sopot; Maria Rell-Bakalarska, Warszawa; Elzbieta Langer-Bieda, Krakow; Jan Brzezicki, Elblag. *Russia:* Marina Stanislav, Moscow; Alexey Maslyanskiy, St. Petersburg; Irina Marusenko, Petrozavodsk; Natalia Shilkina, Yaroslavl; Yurii Shvartz, Saratov; Rimma Kamalova, Ufa; Andrey Rebrov, Saratov; Leysan Myasoutova, Kazan; Elena Zonova, Novosibirsk; Irina Vinogradova, Ulyanovsk. *Singapore:* Tang Ching Lau, Novena; Kam Hon Yoon, Boon Keng. *Ukraine:* Oleksander Dyadyk, Donetsk; Andriy Gnylorybov, Donetsk; Volodymyr Kovalenko, Kiev; Andriy Petrov, Simferopol; Mykola Stanislavchuk, Vinnitsa; Roman Yatsyshyn, Ivano-Frankovsk; Svitlana Smiyan, Ternopil; Vira Tseluyko, Kharkiv; Samvel Turianytsia, Uzhgorod. *USA:* John Budd, St. Louis, Missouri; Mitchell Lowenstein, Palm Harbor, Florida; Michael Miniter, Rock Island, Illinois.

Contributors Study design: SKA, EI, XLX, WR, AG, AB, DB; Data collection, analysis and/or interpretation; drafting or revising the manuscript; approval to submit manuscript: JSS, SKA, EI, XLX, YM, YZ, IN, WR, AG, AB, DB.

Funding This study was funded by Janssen Research & Development, LLC.

Competing interests JSS has received grants from AbbVie, BMS, Janssen, Lilly, MSD, Pfizer and Roche and has served as a consultant for AbbVie, Amgen, Astra-Zeneca, Astro, BMS, GlaxoSmithKline, Janssen, Lilly, MSD, Novartis, Pfizer, Roche, Samsung, Sanofi-Aventis and UCB. SKA served as a Steering Committee member for Janssen. EI served as a trial investigator for Janssen. XLX, YZ, IN, AG, and DB and YM are employees of Janssen Research & Development, LLC, and own stock in Johnson & Johnson, of which Janssen Research & Development, LLC, is a wholly owned subsidiary. AB and WR were employees of Janssen Research & Development, LLC, at the time this work was performed and own stock in Johnson & Johnson, of which Janssen Research & Development, LLC, is a wholly owned subsidiary. WR is currently employed at Sandoz, Inc., Princeton, NJ.

Ethics approval Institutional review board or ethics committee at each site.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Unpublished data from this trial are not currently publicly available.

REFERENCES

1 Emery P, Fleischmann RM, Moreland LW, *et al*. Golimumab, a human anti-tumor necrosis factor α monoclonal antibody, injected subcutaneously every four weeks in methotrexate-naive patients with active rheumatoid arthritis: twenty-four-week results of a phase III, multicenter, randomized, double-blind, placebo-controlled study of golimumab before methotrexate as first-line therapy for early-onset rheumatoid arthritis. *Arthritis Rheum* 2009;60:2272–83.
2 Keystone EC, Genovese MC, Klareskog L, *et al*. Golimumab, a human antibody to tumour necrosis factor α given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: the GO-FORWARD Study. *Ann Rheum Dis* 2009;68:789–96.
3 Klareskog L, van der Heijde D, de Jager JP, *et al*. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. *Lancet* 2004;363:675–81.
4 Maini R, St Clair EW, Breedveld F, *et al*. Infliximab (chimeric anti-tumour necrosis factor α monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. *Lancet* 1999;354:1932–9.
5 Moreland LW, Schiff MH, Baumgartner SW, *et al*. Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial. *Ann Intern Med* 1999;130:478–86.
6 Weinblatt ME, Bingham CO III, Mendelsohn AM, *et al*. Intravenous golimumab is effective in patients with active rheumatoid arthritis despite methotrexate therapy with responses as early as week 2: results of the phase 3, randomised, multicentre, double-blind, placebo-controlled GO-FURTHER trial. *Ann Rheum Dis* 2013;72:381–9.
7 Weinblatt ME, Keystone EC, Furst DE, *et al*. Adalimumab, a fully human anti-tumor necrosis factor α monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis Rheum* 2003;48:35–45.
8 Gabay C, Riek M, Hetland ML, *et al*. Effectiveness of tocilizumab with and without synthetic disease-modifying antirheumatic drugs in rheumatoid arthritis: results from a European collaborative study. *Ann Rheum Dis* 2016;75:1336–42.
9 Iannone F, Gremese E, Atzeni F, *et al*. Longterm retention of tumor necrosis factor-α inhibitor therapy in a large Italian cohort of patients with rheumatoid arthritis from the GISEA registry: an appraisal of predictors. *J Rheumatol* 2012;39:1179–84.
10 Kim W, Min S, Cho M, *et al*. The role of IL-12 in inflammatory activity of patients with rheumatoid arthritis (RA). *Clin Exp Immunol* 2000;119:175–81.
11 Steiner G, Tohidast-Akrad M, Witzmann G, *et al*. Cytokine production by synovial T cells in rheumatoid arthritis. *Rheumatology (Oxford)* 1999;38:202–13.
12 Karonitsch T, von Dalwigk K, Steiner CW, *et al*. Interferon signals and monocytic sensitization of the interferon-γ signaling pathway in the peripheral blood of patients with rheumatoid arthritis. *Arthritis Rheum* 2012;64:400–8.
13 Lubberts E. The IL-23-IL-17 axis in inflammatory arthritis. *Nat Rev Rheumatol* 2015;11:415–29.
14 Miossec P, Korn T, Kuchroo VK. Interleukin-17 and type 17 helper T cells. *N Engl J Med* 2009;361:888–98.
15 Toussirot E. The IL23/Th17 pathway as a therapeutic target in chronic inflammatory diseases. *Inflamm Allergy Drug Targets* 2012;11:159–68.
16 Andersen T, Hvid M, Johansen C, *et al*. Interleukin-23 in early disease development in rheumatoid arthritis. *Scand J Rheumatol* 2015;44:438–42.
17 Kageyama Y, Kobayashi H, Kato N. Infliximab treatment reduces the serum levels of interleukin-23 in patients with rheumatoid arthritis. *Mod Rheumatol* 2009;19:657–62.
18 Kageyama Y, Ichikawa T, Nagafusa T, *et al*. Etanercept reduces the serum levels of interleukin-23 and macrophage inflammatory protein-3 alpha in patients with rheumatoid arthritis. *Rheumatol Int* 2007;28:137–43.
19 Kim HR, Cho ML, Kim KW, *et al*. Up-regulation of IL-23p19 expression in rheumatoid arthritis synovial fibroblasts by IL-17 through PI3-kinase-, NF-κB- and p38 MAPK-dependent signalling pathways. *Rheumatology (Oxford)* 2007;46:57–64.
20 Chabaud M, Durand JM, Buchs N, *et al*. Human interleukin-17: A T cell-derived proinflammatory cytokine produced by the rheumatoid synovium. *Arthritis Rheum* 1999;42:963–70.
21 Leipe J, Grunke M, Dechant C, *et al*. Role of Th17 cells in human autoimmune arthritis. *Arthritis Rheum* 2010;62:2876–85.
22 Kirkham BW, Kavanaugh A, Reich K. Interleukin-17A: a unique pathway in immune-mediated diseases: psoriasis, psoriatic arthritis and rheumatoid arthritis. *Immunology* 2014;141:133–42.
23 Leonardi CL, Kimball AB, Papp KA, *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.
24 Papp KA, Langley RG, Lebwohl M, *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.
25 Kavanaugh A, Puig L, Gottlieb AB, *et al*. Maintenance of clinical efficacy and radiographic benefit through two years of ustekinumab therapy in patients with active psoriatic arthritis: Results from a randomized, placebo-controlled phase III trial. *Arthritis Care Res (Hoboken)* 2015;67:1739–49.
26 McInnes IB, Kavanaugh A, Gottlieb AB, *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.
27 Gordon KB, Duffin KC, Bissonnette R, *et al*. A phase 2 trial of guselkumab versus adalimumab for plaque psoriasis. *N Engl J Med* 2015;373:136–44.
28 Arnett FC, Edworthy SM, Bloch DA, *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.

29 Felson DT, Anderson JJ, Boers M, *et al*. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38:727–35.
30 Prevoo MLL, van't Hof MA, Kuper HH, *et al*. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.
31 Aletaha D, Smolen JS. The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. *Best Pract Res Clin Rheumatol* 2007;21:663–75.
32 Fries JF, Spitz P, Kraines RG, *et al*. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
33 Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). Conceptual framework and item selection. *Medical Care* 1992;30:473–83.
34 Weinblatt ME, Fleischmann R, Huizinga TW, *et al*. Efficacy and safety of certolizumab pegol in a broad population of patients with active rheumatoid arthritis: results from the REALISTIC phase IIIb study. *Rheumatology (Oxford)* 2012;51:2204–14.
35 Genovese MC, Fleischmann R, Kivitz AJ, *et al*. Sarilumab plus methotrexate in patients with active rheumatoid arthritis and inadequate response to methotrexate: results of a phase III study. *Rheumatology (Oxford)* 2015;67:1424–37.
36 Huizinga TW, Fleischmann RM, Jasson M, *et al*. Sarilumab, a fully human monoclonal antibody against IL-6Rα in patients with rheumatoid arthritis and an inadequate response to methotrexate: efficacy and safety results from the randomised SARIL-RA-MOBILITY Part A trial. *Ann Rheum Dis* 2014;73:1626–34.
37 Genovese MC, Durez P, Richards HB, *et al*. Efficacy and safety of secukinumab in patients with rheumatoid arthritis: a phase II, dose-finding, double-blind, randomised, placebo controlled study. *Ann Rheum Dis* 2013;72:863–9.
38 Genovese MC, Greenwald M, Cho CS, *et al*. A phase II randomized study of subcutaneous ixekizumab, an anti-interleukin-17 monoclonal antibody, in rheumatoid arthritis patients who were naive to biologic agents or had an inadequate response to tumor necrosis factor inhibitors. *Ann Rheum Dis* 2014;66:1693–704.
39 Pavelka K, Chon Y, Newmark R, *et al*. A study to evaluate the safety, tolerability, and efficacy of brodalumab in subjects with rheumatoid arthritis and an inadequate response to methotrexate. *J Rheumatol* 2015;42:912–19.
40 Gottlieb AB, Cooper KD, McCormick TS, *et al*. A phase 1, double-blind, placebo-controlled study evaluating single subcutaneous administrations of a human interleukin-12/23 monoclonal antibody in subjects with plaque psoriasis. *Curr Med Res Opin* 2007;23:1081–92.
41 Kremer JM, Genant HK, Moreland LW, *et al*. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2006;144:865–76.
42 Mease P, Genovese MC, Gladstein G, *et al*. Abatacept in the treatment of patients with psoriatic arthritis: results of a six-month, multicenter, randomized, double-blind, placebo-controlled, phase II trial. *Arthritis Rheum* 2011;63:939–48.
43 Tak PP, van der Lubbe PA, Cauli A, *et al*. Reduction of synovial inflammation after anti-CD4 monoclonal antibody treatment in early rheumatoid arthritis. *Arthritis Rheum* 1995;38:1457–65.
44 Bonelli M, Ferner E, Göschl L, *et al*. Abatacept (CTLA-4IG) treatment reduces the migratory capacity of monocytes in patients with rheumatoid arthritis. *Arthritis Rheum* 2013;65:599–607.
45 Bonelli M, Göoschl L, Blüml S, *et al*. Abatacept (CTLA-4Ig) treatment reduces T cell apoptosis and regulatory T cell suppression in patients with rheumatoid arthritis. *Rheumatology (Oxford)* 2016;55:710–20.
46 Ko HJ, Cho ML, Lee SY, *et al*. CTLA4-Ig modifies dendritic cells from mice with collagen-induced arthritis to increase the CD4+CD25+Foxp3+ regulatory T cell population. *J Autoimmun* 2010;34:111–20.
47 Martin DA, Churchill M, Flores-Suarez L, *et al*. A phase Ib multiple ascending dose study evaluating safety, pharmacokinetics, and early clinical response of brodalumab, a human anti-IL-17R antibody, in methotrexate-resistant rheumatoid arthritis. *Arthritis Res Ther* 2013;15:R164.
48 McInnes IB, Mease PJ, Kirkham B, *et al*. Secukinumab, a human anti-interleukin-17A monoclonal antibody, in patients with psoriatic arthritis (FUTURE 2): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 2015;386:1137–46. .
49 Mease PJ, Genovese MC, Greenwald MW, *et al*. Brodalumab, an anti-IL17RA monoclonal antibody, in psoriatic arthritis. *N Engl J Med* 2014;370:2295–306.
50 Langley RG, Elewski BE, Lebwohl M, *et al*. Secukinumab in plaque psoriasis—results of two phase 3 trials. *N Engl J Med* 2014;371:326–38.
51 Papp KA, Reich K, Paul C, *et al*. A prospective phase III, randomized, double-blind, placebo-controlled study of brodalumab in patients with moderate-to-severe plaque psoriasis. *Br J Dermatol* 2016;175:273–86.

OPEN ACCESS

EXTENDED REPORT

Efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy for the treatment of patients with active rheumatoid arthritis (MONARCH): a randomised, double-blind, parallel-group phase III trial

Gerd R Burmester,[1] Yong Lin,[2] Rahul Patel,[2] Janet van Adelsberg,[3] Erin K Mangan,[3] Neil M H Graham,[3] Hubert van Hoogstraten,[2] Deborah Bauer,[2] Juan Ignacio Vargas,[4] Eun Bong Lee[5]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210310).

[1]Department of Rheumatology and Clinical Immunology, Charité—University Medicine Berlin, Free University and Humboldt University Berlin, Berlin, Germany
[2]Sanofi Genzyme, Bridgewater, New Jersey, USA
[3]Regeneron Pharmaceuticals, Inc, Tarrytown, New York, USA
[4]Quantum Research, Puerto Varas, Chile
[5]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

Correspondence to
Dr Gerd R Burmester, Department of Rheumatology and Clinical Immunology, Charité—University Medicine Berlin, Free University and Humboldt University Berlin, Charitéplatz 1, Berlin 10117, Germany; Gerd.Burmester@charite.de

Received 3 August 2016
Revised 13 October 2016
Accepted 15 October 2016
Published Online First 16 November 2016

To cite: Burmester GR, Lin Y, Patel R, *et al*. *Ann Rheum Dis* 2017;**76**:840–847.

ABSTRACT

Objectives To compare efficacy and safety of sarilumab monotherapy with adalimumab monotherapy in patients with active rheumatoid arthritis (RA) who should not continue treatment with methotrexate (MTX) due to intolerance or inadequate response.

Methods MONARCH was a randomised, active-controlled, double-blind, double-dummy, phase III superiority trial. Patients received sarilumab (200 mg every 2 weeks (q2w)) or adalimumab (40 mg q2w) monotherapy for 24 weeks. The primary end point was change from baseline in 28-joint disease activity score using erythrocyte sedimentation rate (DAS28-ESR) at week 24.

Results Sarilumab was superior to adalimumab in the primary end point of change from baseline in DAS28-ESR (−3.28 vs −2.20; p<0.0001). Sarilumab-treated patients achieved significantly higher American College of Rheumatology 20/50/70 response rates (sarilumab: 71.7%/45.7%/23.4%; adalimumab: 58.4%/29.7%/11.9%; all p≤0.0074) and had significantly greater improvement in Health Assessment Questionnaire-Disability Index (p=0.0037). Importantly, at week 24, more patients receiving sarilumab compared with adalimumab achieved Clinical Disease Activity Index remission (7.1% vs 2.7%; nominal p=0.0468) and low disease activity (41.8% vs 24.9%; nominal p=0.0005, supplemental analysis). Adverse events occurred in 63.6% (adalimumab) and 64.1% (sarilumab) of patients, the most common being neutropenia and injection site reactions (sarilumab) and headache and worsening RA (adalimumab). Incidences of infections (sarilumab: 28.8%; adalimumab: 27.7%) and serious infections (1.1%, both groups) were similar, despite neutropenia differences.

Conclusions Sarilumab monotherapy demonstrated superiority to adalimumab monotherapy by improving the signs and symptoms and physical functions in patients with RA who were unable to continue MTX treatment. The safety profiles of both therapies were consistent with anticipated class effects.

Trial registration number NCT02332590.

INTRODUCTION

Biological disease-modifying antirheumatic drugs (bDMARDs) targeting inflammatory cytokines, such as tumour necrosis factor α (TNF-α) or interleukin 6 (IL-6) via the IL-6 receptor (IL-6R), have expanded the treatment options for patients with rheumatoid arthritis (RA).[1–3] Emerging data have demonstrated that patients with inadequate response to conventional synthetic DMARDs (csDMARDs; eg, methotrexate (MTX)) benefit from early and intensive therapy with the addition of bDMARDS, resulting in better preservation of joint structure and function.[4–9] Yet, nearly one-third of patients with RA use biologics as monotherapy due to MTX intolerance or contraindication.[10–13] In addition, increasing data from real-world clinical practice and prescription drug registries across multiple countries indicate that bDMARDs are frequently used as monotherapy, either at the discretion of the physician or because of patient preference.[13–17] The widespread use of bDMARD monotherapy calls for more comparative data to support the optimal selection of approved bDMARDs in clinical practice.

Therapeutic targeting of the IL-6R has been a major advance in the effective treatment of RA, as IL-6R plays a key role in mediating the underlying disease pathophysiology and clinical manifestations of RA.[18–22] In patients with RA, elevated levels of IL-6 in the serum and synovial fluid tightly associate with synovitis, systemic inflammation, bone metabolism, fatigue and joint destruction.[23]

Sarilumab is a human IgG1 monoclonal antibody that binds specifically to both soluble and membrane-bound IL-6Rs (sIL-6Rα and mIL-6Rα) and has been shown to inhibit IL-6-mediated signalling through these receptors. In two previous phase III trials, sarilumab administered subcutaneously at 150 and 200 mg every 2 weeks (q2w) was effective in several patient populations with RA, including MTX inadequate responders[24] and those with an inadequate response or intolerance to TNF inhibitors.[25] In MTX inadequate responders, the addition of sarilumab inhibited radiographic progression and, in both studies, sarilumab achieved

rapid and sustained improvement in disease activity and improved physical function with a manageable safety and tolerability profile consistent with IL-6R blockade.[24–27]

Adalimumab is a globally approved bDMARD targeting TNF-α that is recommended for use in patients who fail to achieve clinical remission with csDMARDs (including MTX) and is an approved monotherapy for those unable to take csDMARDs because of intolerance or contraindication.[2 28] The objective of the phase III MONARCH trial (NCT02332590) was to compare the efficacy and safety of sarilumab and adalimumab monotherapy in patients with active RA who were unsuitable candidates for continued treatment with MTX due to intolerance or inadequate response. Results from this study address the need for data comparing biological monotherapy performance, to help better define strategies for the choice and optimal sequencing of available therapeutics suited for real-world clinical practice.

METHODS

Study design

MONARCH was a multicentre, randomised, active-controlled, double-blind, double-dummy, phase III superiority trial conducted in 86 study centres in Europe, Israel, Russia, South Africa, South America, South Korea and the USA. The first patient was enrolled on 11 February 2015 and the last patient completed week 24 on 20 January 2016. After 24 weeks, patients had the option to enrol in an open-label extension. Results from the 24-week, double-blind treatment period are presented.

Patients were centrally randomised using an interactive voice response system to receive sarilumab 200 mg q2w plus placebo q2w (n=184) or adalimumab 40 mg q2w plus placebo (n=185) in prefilled matching syringes for subcutaneous administration for 24 weeks. Treatment and matching placebo were provided in kits suitable for double-dummy blinding; investigators did not have access to randomisation information except under exceptional medical circumstances. After week 16, dose escalation to weekly administration of adalimumab or matching placebo in the sarilumab group was permitted for patients who did not achieve ≥20% improvement in tender and swollen joint counts.

The protocol was approved by the appropriate ethics committees/institutional review boards and each patient gave written consent before participation in the study. The study was conducted in compliance with institutional review board regulations, the International Conference on Harmonization Guidelines for Good Clinical Practice and the Declaration of Helsinki.

Patient population

Eligible patients were ≥18 years at baseline and those who fulfilled the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism Classification Criteria for RA[29] and ACR class I–III functional status, based on the 1991 revised criteria.[30] Patients were included if they had active RA, defined as ≥6 of 66 swollen and ≥8 of 68 tender joints and high-sensitivity C reactive protein (CRP) ≥8 mg/L or erythrocyte sedimentation rate (ESR) ≥28 mm/hours and 28-joint disease activity score using ESR (DAS28-ESR) >5.1 assessed between screening and randomisation, with disease duration ≥3 months and were, per investigator judgement, either intolerant of or considered inappropriate candidates for continued treatment with MTX, or inadequate responders if treated with an adequate MTX dose (10–25 mg/week or 6–25 mg/week for patients within Asia-Pacific region) for ≥12 weeks. Patients with prior bDMARD experience were excluded.

Efficacy end points

The primary efficacy end point was change from baseline in DAS28-ESR at week 24. Secondary efficacy end points at week 24 included DAS28-ESR remission (<2.6); the Health Assessment Questionnaire-Disability Index (HAQ-DI); ACR 20% (ACR20), 50% (ACR50) and 70% (ACR70) responses; Medical Outcomes Short Form 36 Health Survey (V.2) (SF-36) physical component summary (PCS) score and mental component summary (MCS) score and Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F). Clinical Disease Activity Index (CDAI), a pre-specified secondary end point, was not part of the hierarchy as it was not consistent with regulatory guidance. For a list of end points, see online supplementary table S1.

Safety

Safety assessments included incidence of treatment-emergent adverse events (AEs), serious AEs (SAEs) reported by investigators, along with measured laboratory tests. AEs were described in the Medical Dictionary for Regulatory Activities (V.18.1) preferred-term level, whereas AEs of special interests were identified using pre-specified search criteria. Antidrug antibody (ADA) positivity at two or more consecutive samplings or the last sample analysed during the AE period was classified as persistent.

Statistical analysis

A sample size of 170 patients per group was needed to provide at least 90% power to demonstrate that sarilumab was superior to adalimumab by at least 0.6 units (a clinically relevant difference[26]) on the DAS28-ESR scale using a SD of 1.7 based on prior trials.[26] Efficacy analyses were conducted in the intent-to-treat (ITT) population, which included all randomised patients, including those who increased the dose frequency of adalimumab or matching placebo. Data collected after permanent treatment discontinuation were excluded. Sensitivity analyses and statistical methods are described in the online supplementary appendix.

RESULTS

Patient demographics, baseline characteristics and disposition

The ITT population consisted of 369 patients (185 in the adalimumab group and 184 in the sarilumab group; figure 1). Baseline characteristics and treatment history were generally balanced between groups (table 1).

Patients in the sarilumab group tended to have lower baseline CRP and longer RA duration compared with patients in the adalimumab group, although DAS28-CRP and DAS28-ESR were comparable between groups. Percentages of MTX non-responders and MTX-intolerant patients were also balanced (table 1). The mean highest weekly prior MTX dose was 16.9 mg/week.

The treatment period was completed by most patients (sarilumab: 90%, adalimumab: 84%), with AEs the most common cause of discontinuation (figure 1). The safety population consisted of all patients who received at least one dose of study medication; this population included 368 patients because one patient was randomised to the adalimumab group in error and did not receive study medication.

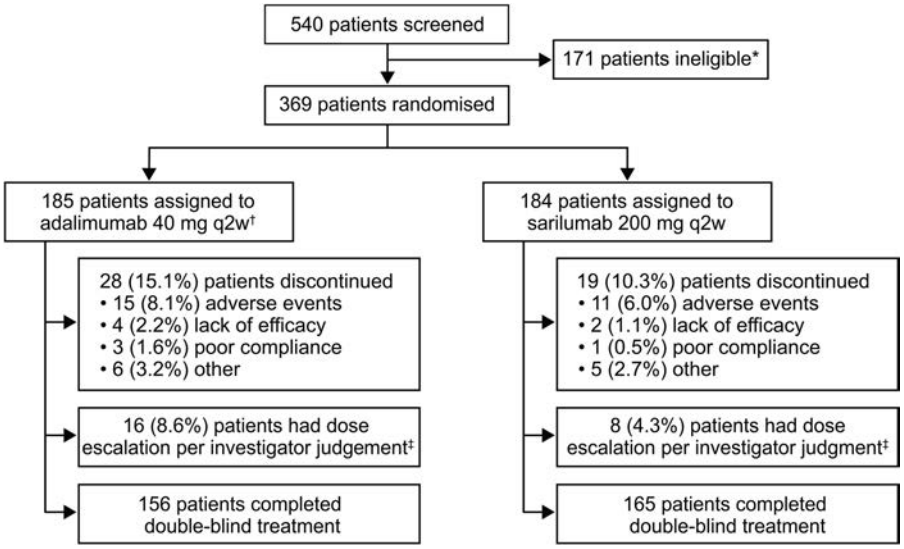


Figure 1 Flow diagram showing patient disposition. *Primary reasons for patient ineligibility were meeting the exclusion criteria related to tuberculosis (12.0%) or failure to meet the inclusion criterion for severity of disease (8.1%). †One patient was randomised but not treated in the adalimumab group. ‡The actual number of patients who received a dose-escalation kit on the basis of meeting protocol criteria were 6 (3.2%) in the adalimumab group and 5 (2.7%) in the sarilumab group. q2w, every 2 weeks.

Efficacy

The primary end point of the study was achieved: sarilumab 200 mg q2w was superior to adalimumab 40 mg q2w in mean change from baseline to week 24 in DAS28-ESR (−3.28 vs −2.20; difference: −1.08; 95% CI −1.36 to −0.79; p<0.0001) (table 2; figure 2A).

Improvements in DAS28-ESR were greater by week 12 in the sarilumab group compared with the adalimumab group (−2.77 vs −1.88; difference: −0.89; 95% CI −1.18 to −0.59; nominal p<0.0001). Compared with adalimumab, the odds of achieving DAS28-ESR remission with sarilumab were approximately three times greater at week 12 (OR: 2.61; 95% CI 1.31 to 5.20; nominal p=0.0051) and approximately five times greater at week 24 (OR: 4.88; 95% CI 2.54 to 9.39; p<0.0001) (figure 3A).

Sensitivity analyses (see online supplementary appendix) for the primary endpoint were consistent with the primary analysis (see online supplementary table S2). Additionally, in a prespecified subgroup analysis, sarilumab demonstrated greater change from baseline in DAS28-ESR at week 24 compared with adalimumab, regardless of previous MTX response (treatment-by-subgroup interaction: intolerant vs inadequate response, p=0.2163; see online supplementary table S3).

Change in DAS28-CRP at week 24 was consistent with DAS28-ESR (−2.86 vs −1.97; difference: −0.88; 95% CI −1.14 to −0.63; nominal p<0.0001). At the first assessment (week 4), mean change in DAS28-CRP was larger in the sarilumab group compared with the adalimumab group (−1.46 vs −1.08; difference: −0.38; 95% CI −0.59 to −0.16; nominal p=0.0005) and the numerical difference between groups continued to increase throughout the study (figure 2B; see online supplementary table S4).

Sarilumab also demonstrated greater efficacy compared with adalimumab in CDAI, a measure of clinical response independent of acute-phase reactants that may favour IL-6 inhibition. Patients receiving sarilumab had a lower mean CDAI score at weeks 12 and 24 compared with patients taking adalimumab (week 24: −28.9 vs −25.2; difference: −3.74; 95% CI −6.02 to −1.47; nominal p=0.0013; see online supplementary table S4). At week 24, more patients receiving sarilumab achieved CDAI remission (13/184 (7.1%) vs 5/185 (2.7%); nominal p=0.0468) and CDAI low disease activity (77/184 (41.8%) vs 46/185 (24.9%); post hoc nominal p=0.0005) compared with those receiving adalimumab (figure 3B; see online supplementary table S4).

The proportion of patients who achieved an ACR20/50/70 response at week 24 was significantly greater in the sarilumab group (71.7%/45.7%/23.4%) than the adalimumab group (58.4%/29.7%/11.9%; all p≤0.0074), with differences observed as early as week 8 (figure 3C). In all three response categories, the between-group difference was >10%. At week 24, both sarilumab and adalimumab had greatly reduced the mean tender (9.0 vs 9.9, out of 68 assessed; p=0.0986) and swollen (4.2 vs 4.8, out of 66 assessed; p=0.0446) joint counts (ACR components described in online supplementary table S4).

The mean improvement in HAQ-DI score from baseline to week 24 was significantly greater in the sarilumab group compared with the adalimumab group (−0.61 vs −0.43; difference: −0.18; 95% CI −0.31 to −0.06; p=0.0037) (table 2). The proportion of patients who demonstrated a clinically meaningful improvement of ≥0.22 units as well as the more stringent ≥0.3 units was higher for patients receiving sarilumab versus those receiving adalimumab (nominal p<0.01 for both) (figure 3D).

At week 24, sarilumab-treated patients had significantly greater improvement in the SF-36 PCS compared with adalimumab-treated patients and improvements were observed as early as week 12 (table 2). Both groups demonstrated similar improvement in SF-36 MCS at week 24. An improvement from baseline to week 24 in FACIT-F score was observed in both groups, with a trend towards greater improvement in the sarilumab group (table 2).

Safety

The incidence of AEs (~64%, both groups) and SAEs (adalimumab, 12 (6.5%) vs sarilumab, 9 (4.9%)) and the rate of discontinuations (adalimumab, 13 (7.1%) vs sarilumab, 11 (6.0%)) were similar between groups (table 3).

One patient in the sarilumab group died of acute cardiac failure secondary to aortic dissection and papillary muscle rupture on day 36.

The incidence of infections was similar between groups (adalimumab, 27.7%; sarilumab, 28.8%). Two patients in each treatment group experienced a serious infection: one mastitis and one infective bursitis with sarilumab and one bacterial arthritis and one upper respiratory tract infection with adalimumab (table 3).

Table 1 Patient demographics and disease characteristics

	Adalimumab 40 mg q2w (n=185)	Sarilumab 200 mg q2w (n=184)
Demographics		
Age, mean±SD, year	53.6±11.9	50.9±12.6
Female, n (%)	150 (81.1)	157 (85.3)
Race, white, n (%)	164 (88.6)	171 (92.9)
Weight, mean±SD, kg	71.8±17.8	72.3±16.5
BMI, mean±SD, kg/m^2	27.3±6.5	27.1±5.6
Geographical region, n (%)*		
Region 1	62 (33.5)	61 (33.2)
Region 2	35 (18.9)	36 (19.6)
Region 3	88 (47.6)	87 (47.3)
Disease and treatment history		
Duration of RA, mean±SD, year	6.6±7.8	8.1±8.1
Rheumatoid factor positive, n (%)†	116 (64.8)	119 (66.9)
Anti-CCP autoantibody positive, n (%)‡	138 (76.7)	134 (75.3)
No. of prior csDMARDs, n (%)		
None	0	0
1	88 (47.6)	83 (45.1)
2	58 (31.4)	57 (31.0)
≥3	39 (21.1)	44 (23.9)
Prior csDMARDs other than MTX, n (%)§		
Sulfasalazine	44 (23.8)	59 (32.1)
Leflunomide	45 (24.3)	42 (22.8)
Hydroxychloroquine	43 (23.2)	41 (22.3)
Prior csDMARDS in combination with MTX, n (%)	44 (23.8)	35 (19.0)
Reason for stopping MTX, n (%)¶		
Inadequate responder	103 (55.7)	97 (52.7)
Intolerant	81 (43.8)	87 (47.3)
Inappropriate for continued treatment	1 (0.5)	0
Concomitant oral corticosteroids, n (%)	104 (56.2)	98 (53.3)
Disease activity, mean±SD		
DAS28-ESR**	6.8±0.8	6.8±0.8
DAS28-CRP**	6.0±0.9	6.0±0.9
Swollen joint count (66 assessed)**	17.5±10.3	18.6±10.7
Tender joint count (68 assessed)**	26.7±13.6	28.0±13.2
CDAI score**	42.4±12.0	43.6±12.1
ESR, mm/h**	47.5±23.2	46.5±21.8
CRP, mg/L**	24.1±31.0	17.4±21.3
HAQ-DI score (0–3)**	1.6±0.6	1.6±0.6
SF-36 physical component score (0–100)††	31.5±6.5	30.8±6.1
FACIT-Fatigue score (0–52)††	24.4±10.3	23.6±8.9
SF-36 mental component score (0–100)††	36.9±11.6	36.4±10.4

*Region 1 (Western countries): Czech Republic, Germany, Hungary, Israel, Spain and USA. Region 2 (South America): Chile and Peru. Region 3 (rest of world): Poland, South Africa, South Korea, Romania, Russia and Ukraine.
†Adalimumab group, n=179; sarilumab group, n=178.
‡Adalimumab group, n=180; sarilumab group, n=178.
§Included if used in >5% of the population.
¶MTX intolerance or inappropriate to continue status was primarily based on clinical judgement of the investigator.
**Higher scores represent more severe disease.
††Lower scores represent more severe disease.
BMI, body mass index; CCP, cyclic citrullinated peptide; CDAI, Clinical Disease Activity Index; CRP, C reactive protein; csDMARD, conventional synthetic disease-modifying antirheumatic drug; DAS28-CRP, 28-joint disease activity score using CRP; DAS28-ESR, DAS28 using erythrocyte sedimentation rate; ESR, erythrocyte sedimentation rate; FACIT, Functional Assessment of Chronic Illness Therapy; HAQ-DI, Health Assessment Questionnaire-Disability Index; MTX, methotrexate; q2w, every 2 weeks; RA, rheumatoid arthritis; SF-36, Medical Outcomes Short Form 36 Health Survey.

One patient in the adalimumab group developed multiple sclerosis. One patient in the sarilumab group was diagnosed with demyelinating polyneuropathy; symptoms began before randomisation. No cases of gastrointestinal perforation, anaphylaxis or lupus-like syndrome were reported in either group.

Injection site reactions were reported in 8 patients (4.3%) in the adalimumab group and 17 patients (9.2%) in the sarilumab

Table 2 Hierarchical order of primary and secondary end points at week 24

	Adalimumab 40 mg q2w (n=185)	Sarilumab 200 mg q2w (n=184)	p Value
Primary end point			
DAS28-ESR			
Mean (SD)	4.5 (1.4)	3.5 (1.4)	
LS mean change from baseline (SE)	−2.20 (0.106)	−3.28 (0.105)	<0.0001
Secondary endpoints			
DAS28-ESR <2.6 (remission), n (%)	13 (7.0)	49 (26.6)	<0.0001
ACR50 response, n (%)	55 (29.7)	84 (45.7)	0.0017
ACR70 response, n (%)	22 (11.9)	43 (23.4)	0.0036
ACR20 response, n (%)	108 (58.4)	132 (71.7)	0.0074
HAQ-DI			
Mean (SD)	1.2 (0.7)	1.0 (0.7)	
LS mean change from baseline (SE)	−0.43 (0.05)	−0.61 (0.05)	0.0037
SF-36 (physical component score)			
LS mean change from baseline (SE)	6.1 (0.6)	8.7 (0.6)	0.0006
FACIT-Fatigue			
LS mean change from baseline (SE)	8.4 (0.7)	10.2 (0.7)	0.0689
SF-36 (mental component score)			
LS mean change from baseline (SE)	6.8 (0.8)	7.9 (0.8)	0.3319

ACR, American College of Rheumatology; DAS28-ESR, 28-joint disease activity score using erythrocyte sedimentation rate; FACIT, Functional Assessment of Chronic Illness Therapy; HAQ-DI, Health Assessment Questionnaire-Disability Index; LS, least squares; q2w, every 2 weeks; SF-36, Medical Outcomes Short Form 36 Health Survey.

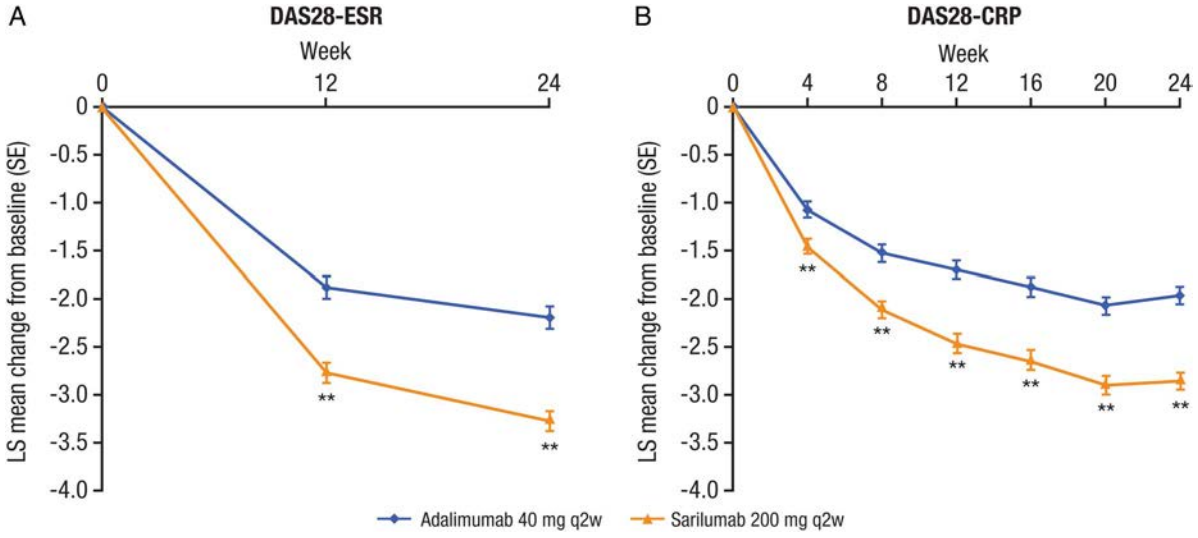


Figure 2 Change from baseline in (A) DAS28-ESR and (B) DAS28-CRP in patients receiving adalimumab 40 mg q2w or sarilumab 200 mg q2w. **p<0.001 versus adalimumab (DAS28-CRP are nominal p values). CRP, C reactive protein; DAS28-ESR, 28-joint disease activity score using erythrocyte sedimentation rate; LS, least squares; q2w, every 2 weeks.

group; two patients in the sarilumab group discontinued as a result. In both groups, the reactions were mild to moderate and the most common AE was erythema.

Neutrophil counts <1.0 G/L occurred more frequently in the sarilumab group compared with the adalimumab group (see online supplementary table S5). Sixteen patients (8.7%) receiving sarilumab and two patients (1.1%) receiving adalimumab had an absolute neutrophil count (ANC) between ≥0.5 and 1 G/L and three patients (1.6%) receiving sarilumab reported an ANC of <0.5 G/L. There was no evidence of an association between decreases in neutrophil counts and risk of infections or serious infections. Infection rates (adalimumab, 51 (27.7%); sarilumab, 53 (28.8%)) were similar between both groups, despite differences in incidence of neutropenia (table 3).

The incidence of alanine aminotransferase (ALT) increases between 1 and 3×upper limit of normal (ULN) was 33.7% in the sarilumab group versus 21.2% in the adalimumab group (see online supplementary table S5). ALT elevations >5×ULN were similar between groups (sarilumab, 1 (0.5%) vs adalimumab, 2 (1.1%)). The mean increase in ALT at week 24 was greater in the sarilumab group (6.1 IU/L) compared with the adalimumab group (2.1 IU/L).

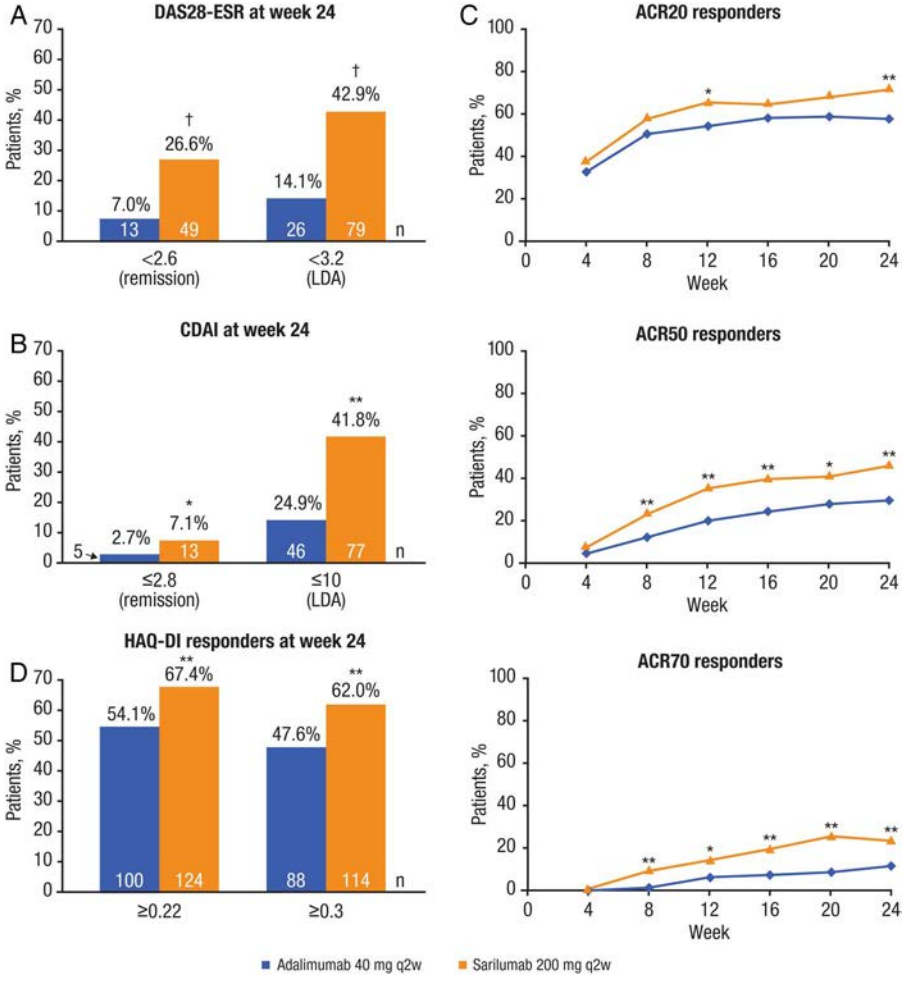


Figure 3 Incidence of (A) DAS28-ESR remission or LDA, (B) ACR20, ACR50 and ACR70 response from weeks 4 to 24, (C) CDAI remission or LDA and (D) HAQ-DI responders achieving ≥0.22 or ≥0.3 units of improvement in patients receiving adalimumab 40 mg q2w or sarilumab 200 mg q2w. *p<0.05 versus adalimumab; **p<0.01 versus adalimumab (CDAI and HAQ-DI responders at week 24 are nominal p values); †p<0.0001 versus adalimumab. ACR, American College of Rheumatology; CDAI, Clinical Disease Activity Index; DAS28-ESR, 28-joint disease activity score using erythrocyte sedimentation rate; HAQ-DI, Health Assessment Questionnaire-Disability Index; LDA, low disease activity; q2w, every 2 weeks.

Reported AEs of serum lipid elevations occurred more frequently in the adalimumab group (8 (4.3%)) than in the sarilumab group (3 (1.6%)) and five patients in the adalimumab group versus two patients in the sarilumab group initiated a lipid-modifying agent during the treatment period. While patients in the sarilumab group demonstrated a greater mean increase from baseline in low-density lipoprotein (LDL) cholesterol compared with patients in the adalimumab group (0.27 mmol/L vs no change), the majority of sarilumab-treated patients did not shift upward in LDL classification (see online supplementary table S6).

ADAs were measured in the sarilumab group; 13 patients tested positive during the AE period. Of these, 5 (2.7%) were defined as persistent ADA because the last sample measured was positive in the ADA assay. No neutralising ADA was detected. The presence of ADA was not associated with hypersensitivity reactions or discontinuations due to lack or loss of efficacy (see online supplementary table S7).

DISCUSSION

Use of biologics as monotherapy is an important therapeutic option for patients with RA when use in combination with MTX or other csDMARDs is unsuitable.[13] In MONARCH, sarilumab was superior to adalimumab in the reduction of disease activity and improvement in the signs and symptoms of RA, as demonstrated by greater reduction in DAS28-ESR. Greater efficacy with sarilumab versus adalimumab was also observed with CDAI, illustrating that the benefits of sarilumab monotherapy extend beyond the pharmacodynamic effects on acute-phase reactants. The odds of CDAI and DAS28 disease remission were greater with sarilumab compared with adalimumab, despite the allowance of adalimumab dose escalation. Additionally, there was no difference in the magnitude of response for patient populations intolerant to MTX versus those with inadequate response, indicating that the robust efficacy outcomes observed with sarilumab were independent of prior MTX use or response.

From the patient's perspective, the most important benefits of RA treatment are to improve functional disability, pain and fatigue.[31–33] Relative to adalimumab, patients receiving sarilumab reported greater improvement in their health status as reflected by differences in SF-36 PCS, HAQ-DI and pain visual analogue scale scores, along with a trend towards greater improvement in fatigue. While the numerical reductions in tender and swollen joint counts were similar between treatment groups, sarilumab-treated patients had less pain and showed greater improvement in global assessments (see online supplementary table S4). These differences reflect that the superiority of objective clinical outcomes observed with sarilumab treatment translate into patient benefits as assessed across a range of patient-reported outcomes.

The safety profiles of sarilumab and adalimumab monotherapy observed in MONARCH were generally comparable. Numerically more patients discontinued with adalimumab because of AEs compared with the sarilumab group. The most common AEs associated with sarilumab were neutropenia and injection site erythema (mostly mild to moderate), while headache and exacerbations of RA were more common in the

Table 3 Safety results

n (%)	Adalimumab 40 mg q2w (n=184)*	Sarilumab 200 mg q2w (n=184)
Overall results		
Patients with any AE	117 (63.6)	118 (64.1)
Patients with any SAE	12 (6.5)	9 (4.9)
Patients with any AE that led to treatment discontinuation	13 (7.1)	11 (6.0)
AEs (≥3% in any treatment group)		
Infections	51 (27.7)	53 (28.8)
Bronchitis	7 (3.8)	12 (6.5)
Nasopharyngitis	14 (7.6)	11 (6.0)
Upper respiratory tract infection	7 (3.8)	3 (1.6)
Neutropenia	1 (0.5)	25 (13.6)
Headache	12 (6.5)	7 (3.8)
Rheumatoid arthritis	7 (3.8)	1 (0.5)
Injection site erythema	6 (3.3)	14 (7.6)
Alanine aminotransferase increased	7 (3.8)	7 (3.8)
Accidental overdose†	11 (6.0)	6 (3.3)
Dyslipidaemia‡	8 (4.3)	3 (1.6)
Serious infections		
Patients with at least one serious infection	2 (1.1)	2 (1.1)
Bursitis, infective	0	1 (0.5)
Mastitis	0	1 (0.5)
Arthritis, bacterial	1 (0.5)	0
Respiratory tract infection	1 (0.5)	0
Deaths§	0	1 (0.5)

*One patient was randomised but not treated in the adalimumab group and was not included in the safety population.
†Protocol defined as ≥2 doses within 11 calendar days or within 6 days for adalimumab-treated patients who switched to weekly dosing.
‡Dyslipidaemia was defined by standardised MedDRA query.
§One patient in the sarilumab group died of acute cardiac failure secondary to aortic dissection and papillary muscle rupture on day 36.
AE, adverse event; MedDRA, Medical Dictionary for Regulatory Activities; q2w, every 2 weeks; SAE, serious adverse event.

adalimumab group. Though neutropenia was more common in the sarilumab group, infection rates were similar between study arms. ADA was monitored in the sarilumab group. There was no relationship between ADA and discontinuations due to lack of efficacy or with hypersensitivity reactions and all instances were non-neutralising.

Overall, the safety and tolerability of sarilumab is consistent across studies[24 25] and comparable with therapeutic targeting of the IL-6 pathway.[26 27] In MONARCH, changes in laboratory values in the sarilumab group, including neutropenia, liver transaminases and total cholesterol, were expected class effects. While the mechanism of neutropenia remains unclear, studies have shown that blockade of IL-6R does not affect neutrophil function.[34] This is consistent with MONARCH and previous sarilumab studies,[24 25] demonstrating that decreased neutrophil counts were not associated with a concurrent increase in infection rate.

In MONARCH, sarilumab monotherapy was associated with lower incidence of ALT elevations compared with previous studies in which sarilumab was administered in combination with csDMARDs.[24 25] Because IL-6 aids in protecting the liver from hepatotoxic agents,[35] IL-6 blockade in combination with MTX, a known hepatotoxicant, may exacerbate the MTX hepatotoxicity observed in some patients.[36 37]

Head-to-head trials comparing the efficacy and safety of two different bDMARD monotherapies in a clinically relevant patient population can aid in defining strategies for optimal patient care. MONARCH expands on results from ADACTA (NCT01119859),[26] showing that blockade of IL-6R is effective in MTX-intolerant patients and in patients with inadequate response to MTX, demonstrating that MTX history does not impact response to therapy. While both MONARCH and ADACTA showed superior efficacy versus adalimumab, MONARCH additionally demonstrated improvement in functional outcomes for patients (table 2). Taken together, the robust MONARCH results further demonstrate that targeting IL-6R may be a preferred treatment option for patients who use biologics as monotherapy.

MONARCH was not without limitations. Although sarilumab plus MTX demonstrated superior radiographic results versus placebo plus MTX,[24] the present study did not evaluate radiographic outcomes after sarilumab monotherapy compared with adalimumab monotherapy. Another limitation of this study is that it did not compare the efficacy of sarilumab monotherapy with sarilumab in combination with MTX. However, as patients intolerant to MTX were the primary intended target population, it would not be feasible to evaluate the addition of sarilumab to MTX in the context of this study.

Collectively, these data demonstrate that sarilumab improves signs and symptoms and functional disability of RA and is an appropriate, effective and superior monotherapy compared with TNF-α inhibition for patients who are unsuitable candidates for continued treatment with MTX due to intolerance or inadequate response.

Acknowledgements Editorial assistance was provided under the direction of the authors by Gretchen Chidester, PhD, and Jennifer Rossi, MA, ELS, MedThink SciCom; Brandy L Bennett, PhD, and Michelle C DeSimone, PhD, DABT, Regeneron Pharmaceuticals and funded by Sanofi Genzyme and Regeneron Pharmaceuticals. The authors thank the patients who participated in this study, the co-investigators for their contribution to the study, staff at the participating centres and the following contributors for providing support: Sophie Bath-Ducrot and Sandra Hankins, for assistance in coordination and execution of the study; Stefano Fiore, for assistance in study design; Petur Wung, for assistance in data acquisition; Alexander Boddy, for assistance in study design, data acquisition and analysis and interpretation of the data and Chieh-I Chen and Clare Proudfoot for assistance in data analysis and interpretation.

Contributors GRB, YL, RP and JvA were involved in the conception and design of the study. GRB, YL, RP, DB, JIV and EBL were involved in data acquisition. GRB, YL, RP, JvA, EKM, NMHG, HvH, DB, JIV and EBL were involved in data analysis and interpretation. GRB, YL and JvA were involved in manuscript drafting. GRB, YL, RP, JvA, EKM, NMHG, HvH, DB, JIV and EBL were involved in critically revising the manuscript and approved the final version.

Funding This study was sponsored by Sanofi Genzyme and Regeneron Pharmaceuticals.

Competing interests GRB has received research grants or consulting fees from AbbVie, Bristol-Myers Squibb, MedImmune, Merck, Pfizer, Roche and UCB and has participated in speakers' bureaus for AbbVie, Bristol-Myers Squibb, Merck, Pfizer, Roche and UCB. EBL has acted as a consultant to Pfizer. JIV has received speaker fees from Roche, Bristol-Myers Squibb and Pfizer and has participated in speakers' bureaus for Bristol-Myers Squibb. YL, RP, HvH and DB are employees of Sanofi Genzyme and may hold stock and/or stock options in the company. JvA, EKM and NMHG are employees of Regeneron Pharmaceuticals and may hold stock and/or stock options in the company.

Ethics approval Individual Study Sites' Institutional Review Boards.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Upchurch KS, Kay J. Evolution of treatment for rheumatoid arthritis. *Rheumatology (Oxford)* 2012;51(Suppl 6):vi28–36.
2 Singh JA, Saag KG, Bridges SL Jr, *et al*. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis Rheumatol* 2016;68:1–26.
3 Smolen JS, Landewé R, Breedveld FC, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Ann Rheum Dis* 2010;69:964–75.
4 Bathon JM, Martin RW, Fleischmann RM, *et al*. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med* 2000;343:1586–93.
5 Keystone EC, Kavanaugh AF, Sharp JT, *et al*. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;50:1400–11.
6 Kremer JM, Blanco R, Brzosko M, *et al*. Tocilizumab inhibits structural joint damage in rheumatoid arthritis patients with inadequate responses to methotrexate: results from the double-blind treatment phase of a randomized placebo-controlled trial of tocilizumab safety and prevention of structural joint damage at one year. *Arthritis Rheum* 2011;63:609–21.
7 Kremer JM, Genant HK, Moreland LW, *et al*. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2006;144:865–76.
8 Lipsky PE, van der Heijde DM, St Clair EW, *et al*., for the Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis With Concomitant Therapy Study Group. Infliximab and methotrexate in the treatment of rheumatoid arthritis. *N Engl J Med* 2000;343:1594–602.
9 Van der Heijde D, van Adelsberg J, Fay J, *et al*. Clinical and radiographic outcomes after 2 years of sarilumab in patients with rheumatoid arthritis [abstract]. *Ann Rheum Dis* 2016;75(Suppl 2):724.
10 Soliman MM, Ashcroft DM, Watson KD, *et al*. Impact of concomitant use of DMARDs on the persistence with anti-TNF therapies in patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. *Ann Rheum Dis* 2011;70:583–9.
11 Listing J, Strangfeld A, Rau R, *et al*. Clinical and functional remission: even though biologics are superior to conventional DMARDs overall success rates remain low—results from RABBIT, the German biologics register. *Arthritis Res Ther* 2006;8:R66.
12 Lee SJ, Chang H, Yazici Y, *et al*. Utilization trends of tumor necrosis factor inhibitors among patients with rheumatoid arthritis in a United States observational cohort study. *J Rheumatol* 2009;36:1611–7.
13 Emery P, Sebba A, Huizinga TW. Biologic and oral disease-modifying antirheumatic drug monotherapy in rheumatoid arthritis. *Ann Rheum Dis* 2013;72:1897–904.
14 Backhaus M, Kaufmann J, Richter C, *et al*. Comparison of tocilizumab and tumour necrosis factor inhibitors in rheumatoid arthritis: a retrospective analysis of 1603 patients managed in routine clinical practice. *Clin Rheumatol* 2015;34:673–81.
15 Pappas DA, Reed GW, Saunders K, *et al*. Characteristics associated with biologic monotherapy use in biologic-naive patients with rheumatoid arthritis in a US registry population. *Rheumatol Ther* 2015;2:85–96.
16 Detert J, Klaus P. Biologic monotherapy in the treatment of rheumatoid arthritis. *Biologics* 2015;9:35–43.
17 Catay E, Bravo M, Rosa J, *et al*. Prevalence of biologics monotherapy in a cohort of patients with rheumatoid arthritis in daily clinical practice. *BMC Musculoskelet Disord* 2016;17:110.
18 Srirangan S, Choy EH. The role of interleukin 6 in the pathophysiology of rheumatoid arthritis. *Ther Adv Musculoskelet Dis* 2010;2:247–56.
19 Steiner G, Tohidast-Akrad M, Witzmann G, *et al*. Cytokine production by synovial T cells in rheumatoid arthritis. *Rheumatology (Oxford)* 1999;38:202–13.
20 Matsumoto T, Tsurumoto T, Shindo H. Interleukin-6 levels in synovial fluids of patients with rheumatoid arthritis correlated with the infiltration of inflammatory cells in synovial membrane. *Rheumatol Int* 2006;26:1096–100.
21 Sugiyama E, Kuroda A, Hori F, *et al*. Serum interleukin-6 level is a sensitive parameter of disease activity in rheumatoid arthritis. *J Clin Rheumatol* 1995;1:93–8.
22 Houssiau FA, Devogelaer JP, Van Damme J, *et al*. Interleukin-6 in synovial fluid and serum of patients with rheumatoid arthritis and other inflammatory arthritides. *Arthritis Rheum* 1988;31:784–8.
23 Dayer JM, Choy E. Therapeutic targets in rheumatoid arthritis: the interleukin-6 receptor. *Rheumatology (Oxford)* 2010;49:15–24.
24 Genovese MC, Fleischmann R, Kivitz AJ, *et al*. Sarilumab plus methotrexate in patients with active rheumatoid arthritis and inadequate response to methotrexate: results of a phase III study. *Arthritis Rheumatol* 2015;67:1424–37.
25 Fleischmann R, van Adelsberg J, Lin Y, *et al*. Sarilumab and non-biologic disease-modifying antirheumatic drugs in patients with active RA and inadequate response or intolerance to TNF inhibitors. *Arthritis Rheumatol*. In press.
26 Gabay C, Emery P, van Vollenhoven R, *et al*. Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (ADACTA): a randomised, double-blind, controlled phase 4 trial. *Lancet* 2013;381:1541–50.
27 Emery P, Rondon J, Garg A, *et al*. Safety and tolerability of subcutaneous sarilumab and intravenous tocilizumab in patients with RA [abstract]. *Arthritis Rheumatol* 2015;67(Suppl 10):971.
28 Humira [package insert]. North Chicago, IL: AbbVie Inc, 2016.
29 Aletaha D, Neogi T, Silman AJ, *et al*. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69:1580–8.
30 Hochberg MC, Chang RW, Dwosh I, *et al*. The American College of Rheumatology 1991 revised criteria for the classification of global functional status in rheumatoid arthritis. *Arthritis Rheum* 1992;35:498–502.
31 Gossec L, Dougados M, Rincheval N, *et al*. Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. *Ann Rheum Dis* 2009;68:1680–5.
32 Sanderson T, Morris M, Calnan M, *et al*. Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. *Arthritis Care Res (Hoboken)* 2010;62:647–56.
33 Strand V, Wright GC, Bergman MJ, *et al*. Patient expectations and perceptions of goal-setting strategies for disease management in rheumatoid arthritis. *J Rheumatol* 2015;42:2046–54.
34 Wright HL, Cross AL, Edwards SW, *et al*. Effects of IL-6 and IL-6 blockade on neutrophil function in vitro and in vivo. *Rheumatology (Oxford)* 2014;53:1321–31.
35 Taub R. Hepatoprotection via the IL-6/Stat3 pathway. *J Clin Invest* 2003;112:978–80.
36 Salliot C, van der Heijde D. Long-term safety of methotrexate monotherapy in patients with rheumatoid arthritis: a systematic literature research. *Ann Rheum Dis* 2009;68:1100–4.
37 Curtis JR, Beukelman T, Onofrei A, *et al*. Elevated liver enzyme tests among patients with rheumatoid arthritis or psoriatic arthritis treated with methotrexate and/or leflunomide. *Ann Rheum Dis* 2010;69:43–7.

EXTENDED REPORT

Risk of diabetes mellitus associated with disease-modifying antirheumatic drugs and statins in rheumatoid arthritis

Gulsen Ozen,[1,2] Sofia Pedro,[3] Marie E Holmqvist,[4] Michael Avery,[5] Frederick Wolfe,[3] Kaleb Michaud[1,3]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-209954).

[1]University of Nebraska Medical Center, Omaha, Nebraska, USA
[2]Faculty of Medicine, Marmara University, Istanbul, Turkey
[3]National Data Bank for Rheumatic Diseases, Wichita, Kansas, USA
[4]Clinical Epidemiology Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden
[5]Regional Health, Rapid City, South Dakota, USA

Correspondence to
Dr Kaleb Michaud, 986270 Nebraska Medical Center, Omaha, NE 68198, USA; kmichaud@unmc.edu

The results have been presented before in abstract form in American College of Rheumatology Annual Meeting 2016.

Received 24 May 2016
Revised 19 October 2016
Accepted 21 October 2016
Published Online First 11 November 2016

ABSTRACT

Objective To investigate the rate of incident diabetes mellitus (DM) in patients with rheumatoid arthritis (RA) and the impact of disease-modifying antirheumatic drug (DMARD) and statin treatments.

Methods We studied patients with RA and ≥1 year participation in the National Data Bank for Rheumatic Diseases without baseline DM from 2000 through 2014. DM was determined by self-report or initiating DM medication. DMARDs were categorised into four mutually exclusive groups: (1) methotrexate monotherapy (reference); (2) any abatacept with or without synthetic DMARDs (3) any other DMARDs with methotrexate; (4) all other DMARDs without methotrexate; along with separate statin, glucocorticoid and hydroxychloroquine (yes/no) variables. Time-varying Cox proportional hazard models were used to adjust for age, sex, socioeconomic status, comorbidities, body mass index and RA severity measures.

Results During a median (IQR) 4.6 (2.5–8.8) years of follow-up in 13 669 patients with RA, 1139 incident DM cases were observed. The standardised incidence ratio (95% CI) of DM in patients with RA (1.37, (1.29 to 1.45)) was increased compared with US adult population. Adjusted HR (95% CI) for DM were 0.67 (0.57 to 0.80) for hydroxychloroquine, 0.52 (0.31 to 0.89) for abatacept (compared with methotrexate monotherapy), 1.31 (1.15 to 1.49) for glucocorticoids and 1.56 (1.36 to 1.78) for statins. Other synthetic/biological DMARDs were not associated with any risk change. Concomitant use of glucocorticoids did not alter DM risk reduction with hydroxychloroquine (HR 0.69 (0.51 to 0.93)).

Conclusions In RA, incidence of DM is increased. Hydroxychloroquine and abatacept were associated with decreased risk of DM, and glucocorticoids and statins with increased risk.

INTRODUCTION

Rheumatoid arthritis (RA) is associated with increased cardiovascular (CV) morbidity and mortality[1–3] likely due to complex interactions between RA-related inflammatory activity, medications and traditional CV disease (CVD) risk factors.[4–6] Among these risk factors, type 2 diabetes mellitus (DM) is one of the most important.[7 8] Although studies investigating DM prevalence in RA have had inconsistent results,[9–12] it has been shown that both inflammatory activity and some RA medications impact glucose metabolism, insulin resistance and consequently DM development.[13–16] Key inflammatory cytokines in RA, particularly tumour necrosis factor-α (TNF-α) and interleukin-6 (IL-6), have been associated with increased adiposity and insulin resistance by triggering key steps in the insulin signalling pathways.[13 15 16]

Besides affecting inflammation, disease-modifying antirheumatic drugs (DMARDs) may also affect DM risk by directly altering glucose metabolism. However, DM risk modification by DMARDs varies in effect.[14 17–20] For example, glucocorticoids (GC), despite strong anti-inflammatory actions, may lead to hyperglycaemia and insulin resistance in a dose-dependent and duration-dependent manner.[14] Alternatively, hydroxychloroquine (HCQ),[21] methotrexate (MTX)[18] and TNF inhibitors (TNFi)[22] have been shown to favourably alter glucose metabolism. Large epidemiological studies also showed decreased risk of new-onset DM with HCQ in RA.[19 20 23] However, these studies had heterogeneous comparison groups and did not address HCQ duration, dose and cessation for DM risk. Similar issues affected studies showing DM risk reduction in TNFi, and none have examined newer biologics.[17 19]

In addition to well-known risk factors such as obesity and physical inactivity, recent meta-analyses and observational studies found that statin use is also associated with increased DM risk in the general population.[24–27] It is unknown whether statins exert the same effect in RA considering its association with decreased RA activity.[28 29]

Given the changing use of treatments and CVD concerns, we sought to investigate the associations of DMARDs, GC and statins with incident DM in patients with RA in a large US-wide observational cohort. We also sought to determine the timing, dosing and sustainability of the observed effects of HCQ in DM risk in patients with RA.

PATIENTS AND METHODS

In a US-wide longitudinal observational cohort, the National Data Bank for Rheumatic Diseases (NDB), participants were mainly recruited from rheumatologists who confirmed the diagnoses. Participants completed semiannual, comprehensive questionnaires as previously described.[30] We included patients with RA who completed at least two questionnaires between January 2000 and December 2014, while those with prevalent DM at study entry were excluded. Follow-up continued

To cite: Ozen G, Pedro S, Holmqvist ME, *et al*. *Ann Rheum Dis* 2017;**76**:848–854.

until the participant reported DM or was censored at death, loss to follow-up or end of study period.

The primary outcome was incident DM defined as patient report of new DM diagnosis or initiating use of an antidiabetic medication. Diagnosis date of DM was assigned a random month of onset during the 6-month period, as done previously.[14]

Treatment exposure was measured at enrolment and every 6 months with questionnaires.[14] Initially, we determined the impact of each DMARD individually on DM risk compared with its non-use (see online supplementary text). From this, we defined four mutually exclusive, hierarchical DMARD groups: (1) MTX monotherapy (reference), (2) any abatacept (ABA) independent of other DMARDs, (3) any other biological (non-ABA) or non-biological DMARDs in combination with MTX, (4) all others. HCQ, GC, statins and non-steroidal anti-inflammatory drugs (NSAIDs) were evaluated separately.

Statistical analysis

Baseline characteristics of the patients with RA with and without incident DM were compared. Crude incidence rates were calculated by dividing the number of events during follow-up by the corresponding person-time at risk. Standardised incidence ratios (SIRs) were estimated using the incidence rates of DM for the United States, reported by the Centers for Disease Control and Prevention, stratified by calendar year, sex and age.[31] Since only incidence rates of DM are reported for individuals under 80 years of age, we excluded those older than 80 at NDB entry when estimating SIRs.

We examined the association between drug exposure and incident DM by using multivariable Cox proportional hazards. A goodness of model fit was compared for each DMARD group. The lowest Akaike information criterion[32] by far was with the above-hierarchical DMARD groups with additional dichotomous HCQ, GC, NSAIDs and statin variables (see online supplementary text).

The final model also included the following: age, sex, ethnicity, income and employment status, Rheumatic Disease Comorbidity Index,[33] hypertension, RA duration (log-transformed), health assessment questionnaire (HAQ), smoking status and body mass index (BMI) categorised according to WHO classification.[34] Three-year calendar periods were also evaluated to account for US treatment trends.

In sensitivity analyses, the impacts of HCQ-ever versus HCQ-never use, daily dose and treatment duration, and prednisone daily dose were examined. The association between treatment duration and risk of incident DM was assessed through Poisson regression. DM risk after discontinuation of HCQ was also evaluated with Cox regression models by selecting patients who were on HCQ (new initiators) but discontinued HCQ (at 3rd, 6th and ≥12th months off HCQ) in comparison with HCQ-never-used patients. The HR was estimated for increasing values of off-treatment duration. For comparison, we took the same approach for new users of MTX. The final model including DMARD groups and other covariates was also analysed in patients without history of CVD (ischaemic heart or peripheral arterial disease, cerebrovascular accident or heart failure), that is, statin use for primary prevention. Potential interactions of concomitant use of GC, either with DMARDs or statins, were also tested.

Finally, a marginal structural model was applied to new initiators of HCQ versus MTX to minimise the bias of confounding by indication. Due to the significant reduction in sample size, results were not presented (see online supplementary text).

In all analyses, treatment exposures were time-dependent covariates. In order to prevent bias from removing observations due to missing data, unanswered covariates of completed questionnaires were replaced by using multiple imputation by chained equations to create imputed datasets for analyses[35] (annual income had 4% missing, all other variables had <1% missing). For non-consecutive observations (8%), the last observation was carried forward, and incident DM diagnosis was assigned at the beginning of the first non-missing 6-month period. All p values were two-sided, conducted at a significance level of 0.05. All statistical analyses were performed using Stata V.14.0 (StataCorp, College Station, Texas, USA).

RESULTS

After excluding 1456 (9.6%) prevalent baseline DM cases, our study had 13 669 patients with RA, with a baseline mean age of 58.6±13.4 years and disease duration of 14.4±12.4 years. The baseline characteristics of the patients by future incident DM and DMARD exposure groups are presented in table 1 and online supplementary table S1, respectively. During a median (IQR) follow-up of 4.6 (2.5–8.8) years, 1139 (8.3%) incident DM cases were observed.

The overall incidence rate (95% CI) for DM was 1.59 (1.50 to 1.68) per 100 person-years. The incidence rate of DM in RA found to be increased (age-adjusted and sex-adjusted SIR, 1.37 (1.29 to 1.45)) compared with the incidence rates in US adult population.[31] The incidence rates in females and males were 1.57 (1.47 to 1.67) and 1.68 (1.48 to 1.91) corresponding to age-adjusted SIRs of 1.39 (1.29 to 1.48) and 1.30 (1.13 to 1.49), respectively. The incidence rates and SIRs by disease activity and DMARD groups, HCQ, GC and statins are shown in table 2.

The fully adjusted time-dependent Cox regression models showed a significant DM risk reduction with HCQ (HR 0.67 (0.57 to 0.80)). In comparison with MTX monotherapy, ABA (HR 0.52 (0.31 to 0.89)) was also associated with DM risk reduction in patients with RA (table 3). Alternatively, the risk of incident DM significantly increased with current use of GC (HR 1.31 (1.15 to 1.49)) or statins (HR 1.56 (1.36 to 1.78)). Besides adjustment for differences in statin-exposed and non-exposed patients (see online supplementary table S2), adjusting for other comorbidities separately (thyroid or CVD) did not change the association with statins (data not shown). Other significant factors associated with increased DM risk included lower annual income, non-Caucasian ethnicity, higher BMI and more comorbidity (see online supplementary table S3).

Sensitivity analysis indicated both HCQ doses of <400 mg/day (median (IQR): 200 mg (200–300)) and ≥400 mg/day (median (IQR): 400 mg (400–600)) were associated with DM risk reduction, though higher doses were more prominent (table 4). Poisson regression models revealed that DM risk reduction with HCQ started after 2 years of treatment (Relative risk (RR) 0.76 (0.58 to 1.00)), and continued to decrease with longer duration of >4 years (RR 0.69 (0.59 to 0.81)). In patients initiating HCQ (N=686), DM risk reduction was consistent with those with prior HCQ exposure, ≥3 years of HCQ treatment (RR 0.44 (0.23 to 0.86)). Patients who initiated and then discontinued HCQ (N=342) had a non-significant risk reduction up to 6 months compared with HCQ never-used patients: HR 0.65 (0.21 to 2.0) for ≥3 months, 0.88 (0.28 to 2.75) for ≥6 months and 1.27 (0.31 to 5.10) for ≥1 year off-HCQ. In contrast, DM risk after MTX initiation and discontinuation was consistently high compared with MTX never-used patients (HR range 1.5–2.8 (0.63 to 8.35)).

Table 1 Baseline characteristics, overall and by future incident diabetes*

	Full RA cohort, N=13 669	Future incident diabetes		p Value
		No diabetes, N=12 530	Diabetes, N=1139	
Age, years	58.6 (13.4)	58.5 (13.5)	59.5 (11.8)	0.011
Female, %	80.3	80.4	79.4	0.406
Non-Hispanic Caucasians, %	93.4	93.6	91.5	0.002
Education, years	13.4 (2.2)	13.5 (2.2)	13.1 (2.0)	0.079
Total annual income (US$10 000)	5.1 (3.1)	5.2 (3.2)	4.5 (2.9)	<0.001
Employed, %	38.9	39.6	32.2	<0.001
Any exercise, %	15.7	15.6	16.6	0.378
BMI, kg/m^2	27.8 (6.4)	27.6 (6.3)	30.8 (7.4)	<0.001
BMI in categories, %				<0.001
<18.5 kg/m^2	2.1	2.2	0.8	
18.5–24.9 kg/m^2	35.3	36.6	20.5	
25.0–29.9 kg/m^2	32.2	32.5	29.3	
30.0–39.9 kg/m^2	25.3	24.1	38.7	
≥40 kg/m^2	5.1	4.5	10.7	
RA duration, years	14.5 (12.4)	14.4 (12.4)	15.0 (12.6)	0.122
HAQ (0–3)	1.0 (0.7)	1.0 (0.7)	1.1 (0.7)	<0.001
Rheumatic Disease Comorbidity Index (0–9)	1.5 (1.4)	1.5 (1.4)	1.7 (1.5)	<0.001
Smoking status, %				<0.001
Never	57.8	58.3	52.9	
Past	30.4	29.9	35.7	
Current	11.7	11.8	11.3	
Hypertension-ever, %	43.3	42.3	53.6	<0.001
CV event-ever, %	18.5	17.8	25.5	<0.001
HCQ-ever, %	50.4	50.8	47.1	0.024
MTX-ever, %	71.1	71.5	67.2	0.002
Any TNFi-ever, %	27.0	27.3	24.6	0.051
Any non-TNFi biologics-ever, %	6.1	6.2	5	0.113
Abatacept-ever, %	3.6	3.7	2.3	0.020
Rituximab-ever, %	1.7	1.8	1.6	0.636
Tocilizumab-ever, %	0.4	0.4	0.3	0.565
Glucocorticoids-ever, %	54.9	55.1	53.7	0.370
Statins-ever, %	14.8	14.6	16.7	0.058

*Values are presented as mean±SD, unless indicated otherwise.
BMI, body mass index; CV, cardiovascular; HAQ, health assessment questionnaire; HCQ, hydroxychloroquine; MTX, methotrexate; RA, rheumatoid arthritis; TNFi, tumour necrosis factor-α inhibitor.

When examining concomitant use of GC, either with HCQ, ABA, or statins, DM risk reduction with HCQ remained significant, while that of ABA vanished. Furthermore, the increased risk with statins was potentiated with GC (table 5). Notably, when HCQ and statins were used together, the increased risk with statins disappeared (HR 0.92 (0.68 to 1.25)). After exclusion of patients with prevalent CVD (N=2535), statin use was still associated with an increased risk of incident DM (HR 2.31 (1.86 to 2.87)) (table 5).

DISCUSSION

In this large US-wide observational cohort study, we found HCQ and ABA associated with decreased risk of incident DM, whereas GC and statins were associated with increased risk.

The risks of incident DM in RA with HCQ,[14 19 20 23] TNFi[17 19] and GC[14] have been evaluated in a few different settings. Previous observational studies reported that HCQ-ever use was associated with 38%–71% DM risk reduction compared with never use,[20 23] and current use was associated with 46% risk reduction compared with the use of any non-biological non-MTX DMARD.[19] A recent UK administrative database study indicated ~20% risk reduction with HCQ use versus non-use.[14] We also observed a 33% risk reduction with HCQ and a 34% reduction with >4 years of HCQ compared with non-users. This relatively lower risk reduction with HCQ in the UK and our cohort compared with previous studies may be due to differing demographics, comparison groups and adjustment for different disease, treatment and diabetes-related covariates (NSAIDs, statins, newer biologics, HAQ and BMI). Finally, more prescription of HCQ to patients with greater DM risk after first publication of this effect may also be a reason for the difference, although we did not observe any trends of HCQ use or effect by calendar year.

From the clinical perspective, determining the minimum dose and duration of HCQ for DM risk reduction and the effects after cessation is also important. We found lower daily doses (<400 mg) of HCQ also decreased DM risk after at least 2 years of treatment (63% of HCQ-exposed were on >2 years of treatment, median (IQR) duration=49 (12–134) months). Previously, it was shown that higher cumulative doses, reflecting

Table 2 Crude incidence rates (95% CI) and SIRs (95% CI) of diabetes in rheumatoid arthritis by disease activity* and treatment compared with US population

	No. of DM	Person-years	Incidence rate (95% CI) per 100 person-years	SIR† (95% CI)
All patients	1139	71 668	1.59 (1.50 to 1.68)	1.37 (1.29 to 1.45)
Remission/low disease activity	550	42 236	1.30 (1.20 to 1.42)	1.13 (1.02 to 1.24)
Moderate/high disease activity	589	29 430	2.00 (1.85 to 2.17)	1.93 (1.77 to 2.11)
Any statins	369	14 851	2.48 (2.24 to 2.75)	2.10 (1.89 to 2.34)
Any glucocorticoids	407	20 369	1.99 (1.81 to 2.20)	1.72 (1.56 to 1.91)
Any HCQ	161	15 603	1.03 (0.88 to 1.20)	0.91 (0.78 to 1.07)
DMARD category				
MTX monotherapy	186	12 761	1.46 (1.26 to 1.68)	1.21 (1.04 to 1.42)
Any abatacept	17	1490	1.14 (0.71 to 1.83)	0.96 (0.58 to 1.59)
Any other DMARD with MTX	224	15 270	1.47 (1.29 to 1.67)	1.27 (1.10 to 1.45)
Other or no DMARDs	551	26 541	2.08 (1.91 to 2.26)	1.82 (1.67 to 1.99)

*Disease activity is determined by 'patient activity scale—PAS.' PAS<3.7 was regarded as remission/low disease activity. PAS≥3.7 was regarded as moderate/high disease activity.
†All participants included were of age <80 years.
DM, diabetes mellitus; DMARD, disease-modifying antirheumatic drug; HCQ, hydroxychloroquine; MTX, methotrexate; SIR, standardised incidence ratio.

longer treatment duration, were associated with DM risk reduction;[20 36] however daily protective doses were not addressed. Improvement in insulin sensitivity has been reported to start within 8 weeks of HCQ treatment in healthy individuals,[37 38] although this short-term improvement has not been shown in stable patients with RA.[39] The only study evaluating the association of DM risk with HCQ treatment duration reported that >4 years of treatment was required for risk reduction.[20] The reason for the difference may be due to lower number of patients, and consequently lower incident DM cases in the group with >4 years HCQ treatment in the previous study (384 of 1808 HCQ-ever used, 21%).[20] In this study, we also showed for the first time that DM risk reduction associated with HCQ tends to continue up to 6 months after cessation of HCQ. This finding is important because it indicates that HCQ offers long-lasting beneficial metabolic effects beyond its anti-inflammatory actions. Future studies are warranted to confirm this finding, as the number of patients who initiated and discontinued HCQ was substantially lower (N=342).

The mechanisms underlying the DM risk reduction with HCQ are usually explained by improvement in insulin sensitivity and pancreatic β-cell functions[37 38] which may be independent of anti-inflammatory actions. The improvement of adiponectin levels without significant change in serum inflammatory cytokines (TNF-α, IL-6) with HCQ in non-diabetic individuals[38] and less significant DM risk reduction with more potent DMARDs support this hypothesis.[19]

Concerning other DMARDs, MTX was associated with a slight decrease in glycosylated haemoglobin in diabetic patients with RA.[18] However, a large cohort study revealed no DM risk reduction with MTX compared with other non-biological DMARDs.[19] We observed a decreased DM risk with MTX when only compared with DMARD non-users in multivariate analysis. As the gold standard therapy for RA, 'MTX monotherapy' was chosen as the reference in our primary analysis. In comparison with 'MTX monotherapy', TNFi monotherapy or concomitant with MTX/other DMARDs also did not modify DM risk. This was inconsistent with prior studies in which TNFi either improved insulin sensitivity[22] or decreased incident DM risk.[17 19] However, these latter studies had relatively younger cohorts, shorter follow-up or no information about BMI or RA activity/severity measures in addition to using 'other DMARDs' as a reference group.[17 19] Finally, the TNFi-associated body fat increase, regardless of disease activity change,[40] may impact DM risk, but this hypothesis necessitates further investigation.

Notably, we also found significant DM risk reduction with ABA compared with MTX monotherapy, which has not been previously reported. Immunologically, ABA may slow the decline of pancreatic β-cell functions in type 1 DM,[41] though its effects on type 2 DM are unknown. Our results, along with evidence from a recent study reporting improvement in insulin sensitivity with 6 months ABA treatment in 15 patients with RA,[42] suggest ABA has favourable effects on insulin resistance. Nevertheless, considering the relatively lower number of patients taking ABA (N=839), investigation in a larger sample is needed.

Another noteworthy finding was the DM risk increase with statins. This association has been previously reported in meta-analyses of randomised controlled trials (RCT) and observational cohort studies of the general population, ranging from 9% to 87% increase.[24–26 43] Although statins were associated with decreased overall mortality,[44] improvement in endothelial functions and atherosclerotic plaques[45 46] and even amelioration in disease activity in RA,[28 29] the impact of statins on the risk of incident DM has not been evaluated before. The effects of statins in primary prevention of CVD in patients with RA have been recently investigated in an RCT of 2986 patients with RA.[47] Despite the early termination of study due to low event rates, a non-significant decrease in CV events was reported in the preliminary results without any DM data.[47] We found a 55% increase in DM risk with statins that was comparable with GC. Adjustment for BMI, smoking, physical activity and other comorbidities including CVD, considering the common risk factors for DM and hypercholesteremia, did not change this association. Moreover in the analysis of statin-using patients for primary prevention of CVD only, the increased risk with statins persisted. Currently, the mechanisms for the higher incidence of DM with statins are not fully understood. The suggested explanations include statin-induced insulin resistance in muscles and liver[48] and genetic variations in 3-hydroxy-3-methylglutaryl-CoA reductase gene.[49] Statin-induced insulin resistance may also be potentiated by chronic inflammation and concurrent GC treatment in RA. To better understand the net effects of statins in RA, including whether CV morbidity and mortality benefits outweigh the risk of DM, further research is warranted.

The well-recognised increased DM risk with GC is also confirmed in this study.[14 50] Additionally, the effects of concomitant

Table 3 Association of different treatments with incident diabetes in patients with RA

Time-dependent treatment variables	Unadjusted HR (95% CI)	p Value	Adjusted HR* (95% CI)	p Value
Statins	1.73 (1.52 to 1.97)	<0.001	1.56 (1.36 to 1.78)	<0.001
Glucocorticoids	1.43 (1.26 to 1.61)	<0.001	1.31 (1.15 to 1.49)	<0.001
HCQ	0.66 (0.55 to 0.78)	<0.001	0.67 (0.57 to 0.80)	<0.001
DMARD groups				
MTX monotherapy (referent)	1.0	–	1.0	–
Any abatacept	0.82 (0.52 to 1.29)	0.39	0.52 (0.31 to 0.89)	0.017
Any other DMARD with MTX	0.98 (0.82 to 1.18)	0.88	0.87 (0.72 to 1.05)	0.158
Other or no DMARDs	1.36 (1.17 to 1.58)	<0.001	1.11 (0.95 to 1.31)	0.190

*Adjusted for age, sex, disease duration, socioeconomic status (employment and income), ethnicity, smoking, hypertension, comorbidity index, BMI, HAQ, NSAID usage and year of entry.
BMI, body mass index; DMARD, disease-modifying antirheumatic drug; HAQ, health assessment questionnaire; HCQ, hydroxychloroquine; MTX, methotrexate; NSAID, non-steroidal anti-inflammatory drug; RA, rheumatoid arthritis.

Table 4 Risk of incident diabetes by prednisone dose, HCQ dose, and duration and concomitant glucocorticoid use

	No. of events/No. of exposure	Adjusted HRs (95% CI)	p Value
Prednisone daily dose, mg/day			
Prednisone non-users (reference)	783/11 686	1.0	–
Prednisone ≤10 mg/day	310/6177	1.27 (1.11 to 1.46)	0.001
Prednisone >10 mg/day	46/1520	1.66 (1.22 to 2.24)	0.001
Concomitant glucocorticoid treatment			
With HCQ	48/2078	0.69 (0.51 to 0.93)	0.014
With abatacept	8/427	0.60 (0.28 to 1.29)	0.188
With statins	127/1724	2.03 (1.65 to 2.50)	<0.001
HCQ daily dose, mg/day			
HCQ non-user (reference)	832/11 753	1.0	–
HCQ <400 mg/day	44/1938	0.71 (0.52 to 0.96)	0.025
HCQ ≥400 mg/day	117/3751	0.66 (0.55 to 0.81)	<0.001
HCQ treatment duration		**Relative risk† (95% CI)**	
HCQ-never use (referent)	517/5631	1.0	–
HCQ-ever use (regardless of duration)	584/7480	0.84 (0.74 to 0.95)	0.006
HCQ: ≤1 year	195/2038	0.93 (0.79 to 1.10)	0.501
HCQ: 1–2 years	74/693	0.96 (0.76 to 1.21)	0.226
HCQ: 2–4 years	63/983	0.76 (0.58 to 1.00)	0.049
HCQ: >4 years	252/3766	0.69 (0.59 to 0.81)	<0.001

*Adjusted for age, sex, disease duration, socioeconomic status (employment and income), ethnicity, smoking, hypertension, comorbidity index, BMI, HAQ score, NSAID usage and year of entry.
†Multivariable Poisson regression analysis stratified for the same covariates given above.
BMI, body mass index; HAQ, health assessment questionnaire; HCQ, hydroxychloroquine; NSAID, non-steroidal anti-inflammatory drug.

GC with other DMARDs were examined. Persistent DM risk reduction with GC and HCQ may indicate a therapeutic approach to decrease DM risk in patients who require long-term GC treatment. This effect was also seen for concomitant use of statins and HCQ. Given that both drugs are commonly used in RA, further investigations are needed to clarify whether GC affects lipid-lowering actions and CV mortality impact of statins.

Our study has important limitations to address. First, although we found a decreased DM risk with HCQ and increased DM risk with statins, the preferential prescription of HCQ to patients with less severe disease and statins to patients who may already be at high risk for DM with frequent presence of obesity, physical inactivity and family history of lipid/glucose metabolism disorders may lead to confounding by treatment indication. To minimise this, we adjusted for several confounding factors related to disease activity/severity and DM risk. We also applied marginal structural models to new initiators of HCQ or MTX (data not shown) and found a non-significant risk reduction with HCQ compared with MTX. With this methodology, less-biased inferences were obtained for the effect of time-varying treatment in the presence of time-varying confounders, which were simultaneously affected by earlier treatment and will affect later treatment. However, restricting the study to new users, with the goal of approximating a RCT, drastically reduced the sample size and study power (only 9% of patients taking HCQ were new users). Second, serological status, tender/swollen joint counts and acute phase reactants were unavailable; therefore, patient activity scale and HAQ measures were used to adjust for RA activity/severity, and the best performing model included HAQ. Finally, we did not have access to plasma glucose or HbA1c values, and could not determine with great accuracy as to when patients developed DM, only when they were diagnosed, which could impact treatment associations.

In conclusion, our findings suggest that DM incidence is increased in patients with RA and that HCQ and ABA are

Table 5 Association between DMARDs and statins and incident diabetes in patients without cardiovascular disease (N=11,134; incident diabetes N=837 (7.8%))

	Adjusted HR* (95% CI)	p Value
Statins	2.31 (1.86 to 2.87)	<0.001
Glucocorticoids	1.50 (1.22 to 1.85)	<0.001
HCQ	0.62 (0.47 to 0.81)	0.001
DMARD groups		
MTX monotherapy (referent)	1.0	–
Any abatacept	0.24 (0.10 to 0.76)	0.015
Any other DMARD with MTX	0.83 (0.62 to 1.10)	0.203
Other or no DMARDs	1.08 (0.84 to 2.88)	0.561

*Adjusted for age, sex, disease duration, socioeconomic status (employment and income), ethnicity, smoking, hypertension, comorbidity index, BMI, HAQ score, NSAID usage and year of entry.
BMI, body mass index; DMARDs, disease-modifying antirheumatic drugs; HAQ, health assessment questionnaire; HCQ, hydroxychloroquine; MTX, methotrexate; NSAID, non-steroidal anti-inflammatory drug.

associated with reduced risk of incident DM in RA while GC and statins are associated with increased risk. HCQ confers a sustainable and treatment duration-dependent favourable effect and eliminates the increased risk associated with GC or statins. Considering the increased CV mortality in RA and the importance of DM to this outcome, our findings can inform clinicians about determining the appropriate treatment decisions in high DM-risk patients with RA. Although further research is required to better understand the effects of statins on RA, given the more frequent presence of other CV risk factors in statin-using patients, careful monitoring for DM should be considered in these patients.

Contributors All authors participated in the conception and design of the study. GO, SP and KM analysed and drafted the manuscript. All authors contributed to the interpretation of the results and reviewed the manuscript.

Funding KM was supported by the Rheumatology Research Foundation Investigator Award.

Competing interests None declared.

Ethics approval This study was conducted with the approval of the Via Christi Regional Medical Center Institutional Review Board, and patients provided informed written consent prior to study enrolment.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Avina-Zubieta JA, Thomas J, Sadatsafavi M, *et al*. Risk of incident cardiovascular events in patients with rheumatoid arthritis: a meta-analysis of observational studies. *Ann Rheum Dis* 2012;71:1524–9.
2 Meune C, Touzé E, Trinquart L, *et al*. Trends in cardiovascular mortality in patients with rheumatoid arthritis over 50 years: a systematic review and meta-analysis of cohort studies. *Rheumatology (Oxford)* 2009;48:1309–13.
3 Solomon DH, Karlson EW, Rimm EB, *et al*. Cardiovascular morbidity and mortality in women diagnosed with rheumatoid arthritis. *Circulation* 2003;107:1303–7.
4 Gonzalez A, Maradit Kremers H, Crowson CS, *et al*. Do cardiovascular risk factors confer the same risk for cardiovascular outcomes in rheumatoid arthritis patients as in non-rheumatoid arthritis patients? *Ann Rheum Dis* 2008;67:64–9.
5 Solomon DH, Kremer J, Curtis JR, *et al*. Explaining the cardiovascular risk associated with rheumatoid arthritis: traditional risk factors versus markers of rheumatoid arthritis severity. *Ann Rheum Dis* 2010;69:1920–5.
6 Baghdadi LR, Woodman RJ, Shanahan EM, *et al*. The impact of traditional cardiovascular risk factors on cardiovascular outcomes in patients with rheumatoid arthritis: a systematic review and meta-analysis. *PLoS ONE* 2015;10:e0117952.
7 Dinesh Shah A, Langenberg C, Rapsomaniki E, *et al*. Type 2 diabetes and incidence of a wide range of cardiovascular diseases: a cohort study in 1.9 million people. *Lancet* 2015;385(Suppl 1):S86.
8 Shah AD, Langenberg C, Rapsomaniki E, *et al*. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people. *Lancet Diabet Endocrinol* 2015;3:105–13.
9 Han C, Robinson DW Jr, Hackett MV, *et al*. Cardiovascular disease and risk factors in patients with rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. *J Rheumatol* 2006;33:2167–72.
10 Simard JF, Mittleman MA. Prevalent rheumatoid arthritis and diabetes among NHANES III participants aged 60 and older. *J Rheumatol* 2007;34:469–73.
11 Solomon DH, Curhan GC, Rimm EB, *et al*. Cardiovascular risk factors in women with and without rheumatoid arthritis. *Arthritis Rheum* 2004;50:3444–9.
12 Solomon DH, Love TJ, Canning C, *et al*. Risk of diabetes among patients with rheumatoid arthritis, psoriatic arthritis and psoriasis. *Ann Rheum Dis* 2010;69:2114–17.
13 Chung CP, Oeser A, Solus JF, *et al*. Inflammation-associated insulin resistance: differential effects in rheumatoid arthritis and systemic lupus erythematosus define potential mechanisms. *Arthritis Rheum* 2008;58:2105–12.
14 Movahedi M, Beauchamp ME, Abrahamowicz M, *et al*. Risk of incident diabetes associated with dose and duration of oral glucocorticoid therapy in patients with rheumatoid arthritis. *Arthritis Rheum* 2016;68:1089–98.
15 Schultz O, Oberhauser F, Saech J, *et al*. Effects of inhibition of interleukin-6 signalling on insulin sensitivity and lipoprotein (a) levels in human subjects with rheumatoid diseases. *PLoS ONE* 2010;5:e14328.
16 Tam LS, Tomlinson B, Chu TT, *et al*. Impact of TNF inhibition on insulin resistance and lipids levels in patients with rheumatoid arthritis. *Clin Rheumatol* 2007;26:1495–8.
17 Antohe JL, Bili A, Sartorius JA, *et al*. Diabetes mellitus risk in rheumatoid arthritis: reduced incidence with anti-tumor necrosis factor α therapy. *Arthritis Care Res (Hoboken)* 2012;64:215–21.
18 Rekedal LR, Massarotti E, Garg R, *et al*. Changes in glycosylated hemoglobin after initiation of hydroxychloroquine or methotrexate treatment in diabetes patients with rheumatic diseases. *Arthritis Rheum* 2010;62:3569–73.
19 Solomon DH, Massarotti E, Garg R, *et al*. Association between disease-modifying antirheumatic drugs and diabetes risk in patients with rheumatoid arthritis and psoriasis. *JAMA* 2011;305:2525–31.
20 Wasko MC, Hubert HB, Lingala VB, *et al*. Hydroxychloroquine and risk of diabetes in patients with rheumatoid arthritis. *JAMA* 2007;298:187–93.
21 Penn SK, Kao AH, Schott LL, *et al*. Hydroxychloroquine and glycemia in women with rheumatoid arthritis and systemic lupus erythematosus. *J Rheumatol* 2010;37:1136–42.
22 Gonzalez-Gay MA, De Matias JM, Gonzalez-Juanatey C, *et al*. Anti-tumor necrosis factor-alpha blockade improves insulin resistance in patients with rheumatoid arthritis. *Clin Exp Rheum* 2006;24:83–6.
23 Bili A, Sartorius JA, Kirchner HL, *et al*. Hydroxychloroquine use and decreased risk of diabetes in rheumatoid arthritis patients. *J Clin Rheumatol* 2011;17:115–20.
24 Mansi I, Frei CR, Wang CP, *et al*. Statins and new-onset diabetes mellitus and diabetic complications: a retrospective cohort study of US healthy adults. *J Gen Intern Med* 2015;30:1599–610.
25 Preiss D, Seshasai SR, Welsh P, *et al*. Risk of incident diabetes with intensive-dose compared with moderate-dose statin therapy: a meta-analysis. *JAMA* 2011;305:2556–64.
26 Sattar N, Preiss D, Murray HM, *et al*. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet* 2010;375:735–42.
27 Waters DD, Ho JE, DeMicco DA, *et al*. Predictors of new-onset diabetes in patients treated with atorvastatin: results from 3 large randomized clinical trials. *J Am Coll Cardiol* 2011;57:1535–45.
28 Lv S, Liu Y, Zou Z, *et al*. The impact of statins therapy on disease activity and inflammatory factor in patients with rheumatoid arthritis: a meta-analysis. *Clin Exp Rheumatol* 2015;33:69–76.
29 McCarey DW, McInnes IB, Madhok R, *et al*. Trial of Atorvastatin in Rheumatoid Arthritis (TARA): double-blind, randomised placebo-controlled trial. *Lancet* 2004;363:2015–21.
30 Wolfe F, Michaud K. The National Data Bank for rheumatic diseases: a multi-registry rheumatic disease data bank. *Rheumatology (Oxford)* 2011;50:16–24.
31 Centers for Disease Control and Prevention. Crude and Age-Adjusted Incidence of Diagnosed Diabetes per 1,000 Population Aged 18–79 Years, United States, 1980–2014. 2015. http://www.cdc.gov/diabetes/statistics/incidence/fig2.htm
32 Akaike H. A new look at the statistical model identification. *Autom Control IEEE Trans* 1974;19:716–23.
33 Michaud K, Wolfe F. Comorbidities in rheumatoid arthritis. *Best Prac Res Clin Rheumatol* 2007;21:885–906.
34 Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser* 2000;894:i-xii, 1–253.
35 Rubin DB. *Multiple imputation for nonresponse in surveys*. New York: Wiley, 1987: xxix, 258p.
36 Chen YM, Lin CH, Lan TH, *et al*. Hydroxychloroquine reduces risk of incident diabetes mellitus in lupus patients in a dose-dependent manner: a population-based cohort study. *Rheumatology (Oxford)* 2015;54:1244–9.
37 Mercer E, Rekedal L, Garg R, *et al*. Hydroxychloroquine improves insulin sensitivity in obese non-diabetic individuals. *Arthritis Res Ther* 2012;14:R135.
38 Wasko MC, McClure CK, Kelsey SF, *et al*. Antidiabetogenic effects of hydroxychloroquine on insulin sensitivity and beta cell function: a randomised trial. *Diabetologia* 2015;58:2336–43.

39 Solomon DH, Garg R, Lu B, *et al*. Effect of hydroxychloroquine on insulin sensitivity and lipid parameters in rheumatoid arthritis patients without diabetes mellitus: a randomized, blinded crossover trial. *Arthritis Care Res (Hoboken)* 2014;66:1246–51.

40 Engvall IL, Tengstrand B, Brismar K, *et al*. Infliximab therapy increases body fat mass in early rheumatoid arthritis independently of changes in disease activity and levels of leptin and adiponectin: a randomised study over 21 months. *Arthritis Res Ther* 2010;12:R197.

41 Orban T, Bundy B, Becker DJ, *et al*. Co-stimulation modulation with abatacept in patients with recent-onset type 1 diabetes: a randomised, double-blind, placebo-controlled trial. *Lancet* 2011;378:412–19.

42 Ursini F, Russo E, Letizia Hribal M, *et al*. Abatacept improves whole-body insulin sensitivity in rheumatoid arthritis: an observational study. *Medicine (Baltimore)* 2015;94:e888.

43 Ridker PM, Pradhan A, MacFadyen JG, *et al*. Cardiovascular benefits and diabetes risks of statin therapy in primary prevention: an analysis from the JUPITER trial. *Lancet* 2012;380:565–71.

44 Schoenfeld SR, Lu L, Rai SK, *et al*. Statin use and mortality in rheumatoid arthritis: a general population-based cohort study. *Ann Rheum Dis* 2016;75:1315–20.

45 Ikdahl E, Hisdal J, Rollefstad S, *et al*. Rosuvastatin improves endothelial function in patients with inflammatory joint diseases, longitudinal associations with atherosclerosis and arteriosclerosis: results from the RORA-AS statin intervention study. *Arthritis Res Ther* 2015;17:279.

46 Rollefstad S, Ikdahl E, Hisdal J, *et al*. Rosuvastatin-induced carotid plaque regression in patients with inflammatory joint diseases: the rosuvastatin in rheumatoid arthritis, ankylosing spondylitis and other inflammatory joint diseases study. *Arthritis Rheum* 2015;67:1718–28.

47 Kitas GD, Nightingale P, Armitage J, *et al*. Trial of atorvastatin for the primary prevention of cardiovascular events in patients with rheumatoid arthritis [Abstract]. *Arthritis Rheumatol* 2015;67 (suppl).

48 Mallinson JE, Constantin-Teodosiu D, Sidaway J, *et al*. Blunted Akt/FOXO signalling and activation of genes controlling atrophy and fuel use in statin myopathy. *J Physiol (Lond)* 2009;587:219–30.

49 Swerdlow DI, Preiss D, Kuchenbaecker KB, *et al*. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2015;385:351–61.

50 Wolfe F, Michaud K. The risk of myocardial infarction and pharmacologic and nonpharmacologic myocardial infarction predictors in rheumatoid arthritis: a cohort and nested case-control analysis. *Arthritis Rheum* 2008;58:2612–21.

EXTENDED REPORT

The patient perspective on absence of disease activity in rheumatoid arthritis: a survey to identify key domains of patient-perceived remission

Lilian H D van Tuyl,[1] Martina Sadlonova,[2] Sarah Hewlett,[3] Bev Davis,[4] Caroline Flurey,[3] Niti Goel,[5] Laure Gossec,[6] Cecilie Heegaard Brahe,[7] Catherine L Hill,[8] Wijnanda Hoogland,[9] John Kirwan,[10] Merete L Hetland,[7] Dirkjan van Schaardenburg,[1] Josef S Smolen,[2] Tanja Stamm,[2] Marieke Voshaar,[9] George A Wells,[11] Maarten Boers[1,12]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-209835).

For numbered affiliations see end of article.

Correspondence to
Dr Lilian H D van Tuyl, Department of Rheumatology, VU University Medical Center, 3A56, Amsterdam 1007 MB, The Netherlands; L.vantuyl@vumc.nl

Received 3 May 2016
Revised 25 October 2016
Accepted 5 November 2016
Published Online First 30 November 2016

ABSTRACT

Background Guidelines suggest treatment in rheumatoid arthritis (RA) to target remission, in close consultation with the patient. Our recent qualitative study of the patients' perspective on remission in RA identified 26 domains. The current study aimed to identify a short list of the most important aspects to inform future research.

Methods Patients with RA from the Netherlands, the UK, Austria, Denmark, France and the USA completed a survey that contained all domains identified in our qualitative study. They rated domains for importance ('not important', 'important' or 'essential' to characterise a period of remission) and if important or essential, whether this domain needs to be 'less', 'almost gone' or 'gone' to reflect remission. Respondents were also asked to determine their personal top 3 most important/essential domains. Frequency of specific domains in the top 3 was calculated, and domains were sorted on the percentage of patients that evaluated a particular domain as 'essential'.

Results Of 274 respondents, 75% were female, mean (SD) age 57(13) years, disease duration 12(9) years. The top 3 were as follows: pain (67%), fatigue (33%) and independence (19%); domains most frequently rated as 'essential' were as follows: pain (60%), being mobile (52%), physical function (51%), being independent (47%) and fatigue (41%). Pain needed to be less (13%), almost gone (42%) or gone (45%) to reflect remission. Similar patterns were seen for fatigue, independence, mobility and physical functioning.

Conclusion Patients identified pain, fatigue and independence as the most important domains of RA disease activity that need to be improved to reflect remission.

INTRODUCTION

Patients with rheumatoid arthritis (RA) are at risk of severe bone and cartilage damage in affected joints, causing chronic pain, fatigue and other extra-articular manifestations with a significant impact on daily life. The degree of disease activity and response to treatment are traditionally determined by evaluation of the RA core set or indices derived thereof.[1 2] The core set contains the patient-reported outcomes (PROs) physical function, pain and global assessment of disease activity. Although not consistently associated with joint damage and differential response in disease stages, these PROs have been found to be at least as relevant as more 'objective' physical and biochemical measures in assessing baseline disease status, improvement during interventions or prediction of long-term outcome.[3 4] Moreover, the relevance of some of them, especially those evaluating physical function, is revealed by observations that increasing joint damage causes increasing irreversibility of functional impairment, even if clinical activity has subsided into remission.[5 6]

It is becoming increasingly clear both inside and outside rheumatology that patients are crucial partners in obtaining relevant information, adding unique skills, values and experiences to research.[7] Patients have identified domains such as fatigue and sleep quality to be important and thus core areas for measurement. Subsequent research has shown measurement of fatigue, one of the most important problems identified by patients with RA, to be highly reliable, sensitive to change and an independent determinant of disease activity.[8–10] As a consequence, the scientific community now recognises fatigue as a core PRO to be measured in all RA clinical trials.[11 12] Other products of close cooperation between patients and professionals include the recent development of patient-derived scores to capture the impact of RA and psoriatic arthritis on daily life.[13 14]

In the last decade, the development of new drugs for the treatment of RA has made a state of minimal disease activity and even remission an attainable goal in most patients.[15–18] Because treatments are increasingly targeted at achieving remission, a good definition of remission is vital. In 2011, the three leading international rheumatology organisations, that is, the American College of Rheumatology (ACR), the European League Against Rheumatism (EULAR) and the Outcome Measures in Rheumatology group (OMERACT), led the initiative which redefined remission in RA.[19–21] To this end, all important prognostic factors and outcome measures available in clinical trial data were evaluated for their potential use in defining remission. However, this included only the

To cite: van Tuyl LHD, Sadlonova M, Hewlett S, *et al*. *Ann Rheum Dis* 2017;**76**:855–861.

three core set PROs patient global, pain and physical function, as data on other potential important aspects of remission from the patient perspective were not available.

In response, both patients and professionals identified the need to study the concept of remission from a patient perspective, to evaluate whether additional domains (and PROs) could optimise targeted therapy.[22] Therefore, we recently undertook a qualitative study to understand the patient perspective on remission in RA.[23] Three major themes of patient-perceived remission emerged:

1. Symptoms such as pain, stiffness and fatigue would either be absent or be reduced in intensity.

2. The impact of the disease on daily life would diminish as shown by increased independence, the ability to do valued activities, improved mood and the ability to cope.

3. Remission would lead to a return to normality, including the ability to work, enjoy one's family role and be seen as normal by other people. Patients felt that the concept of remission was influenced by ageing, side effects of medication, comorbidities, accrued damage to joints and disease duration. This qualitative research identified many domains of interest to patients, but did not indicate the importance of one domain over another.

The aim of the current descriptive study was to determine the importance of specific symptoms, aspects of disease impact and normality in defining remission in RA from the patient perspective through a survey, to complete the information necessary for optimal clinical management.

METHODS

Patients

Patients >18 years of age with a confirmed diagnosis of RA (2010 criteria)[24] receiving usual care in one of five centres (VU University Medical Center/Reade in Amsterdam, the Netherlands; Medical University of Vienna, Austria; Bristol Royal Infirmary in Bristol, UK; Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark; Université Pierre et Marie Curie and Hopital Pitie-Salpetriere Paris in Paris) were invited to participate in this study. Medical ethical committees in these centres approved the study protocol where applicable and patients gave their informed consent before participation. In addition, a fixed sample of 50 patients with RA was recruited through a known community (MediGuard.org) with pre-existing consent to contact for research purposes in the USA.[25]

Eligible patients had to speak, read and write the local language sufficiently to understand the study and complete the survey (physician's judgement where applicable).

In addition, all patient representatives who attended OMERACT 12 in 2014 were invited to participate by email,[26] as well as all the patients who participated in one of the nine focus group discussions prior to this study.[23]

Data collection process

In Bristol, Amsterdam, Vienna and Copenhagen, the surveys and reply envelopes were distributed in the clinic; in Paris and the USA, they were distributed by email.

As the word 'remission' is a common word in the English language and might imply certain presumptions, this term was not used during the recruitment and data collection phase. Instead, remission was formulated as 'disease activity as good as gone'. Where available in routine practice, a recent 28-joint count, physician global assessment and acute phase reactant were collected from the hospital files, within a period of 3 months before or after completion of the survey.

The survey

The goal of the survey was to determine a short list of the most important items that reflect remission according to patients with RA. The survey (see online supplementary file I) contained all 26 domains of remission that were identified in previous focus group discussions in Bristol, Vienna and Amsterdam,[23] formulated as items which patients were asked to rate for importance. In addition, patients were asked to add any missing aspects of remission in free text fields. During the qualitative study, patients had indicated that demographic and disease-specific aspects were important to interpret the data, and so information on age, gender, disease duration, comorbidities and accrued joint damage was collected, all in a self-reported manner using the Routine Assessment of Patient Index Data 3 (RAPID3), with remission cut-off defined as 3 or less on a scale of 0 to 30.[27] Where possible, the Clinical Disease Activity Index (CDAI) was calculated, with remission cut-off ≤2.8.[28]

German, French, Danish and Dutch versions of the survey were prepared by translation and back translation by the research team to verify the terminology. The language used was carefully written, based on the focus group terminology, and was reviewed by patient research partners (WH, MV and BD) to ensure that the instructions were clear to patients and that each item was understandable in terms of the RA symptoms and experience.

To reduce any order effect on decision making, two versions of the survey were distributed in the clinics (but not for the emailed assessments) with the domains and the items within them randomly ordered.

Analysis

Patient characteristics were summarised as mean (SD) or median (IQR) where applicable. Fulfilment of RAPID3 remission, ACR/EULAR remission and CDAI remission was compared with the patients' self-reported judgement of remission ('disease as good as gone': no/yes) to determine concordance between the clinical definition of remission and the patient's judgement of remission, quantified using the κ measure for agreement (with 0.75 as excellent, 0.40–0.75 as fair to good and below 0.40 as poor); 2 by 2 tables and χ^2 tests were used.

To determine the importance of domains, first, frequency of a particular domain mentioned in the top 3 was calculated. Second, domains that >30% of patients identified as 'not important' were removed. The remaining domains were sorted on the percentage of patients that evaluated a particular domain as 'essential'.

To evaluate robustness of the results, data were stratified by the influential factors as identified by patients in the qualitative study, including self-reported age (above or below 50 years), gender (male/female), disease duration (more or less than 2 years), comorbidity (no/yes) and accrued joint damage (no/yes), to see if these factors influence the patient perspective on remission. In addition, data were stratified by country and location of filling out the survey (clinic visit or by email). χ^2 tests were used to determine statistical significance (if $p<0.05$) where relevant.

RESULTS

A total of 274 patients completed the questionnaire: 54 from the Netherlands, 33 from the UK, 51 from Austria, 43 from Denmark, 43 from France and 50 from the USA. Response rate in the Netherlands and France was 59% and 42%, respectively.

The population was typical for RA (table 1), with 75% females, mean (SD) age of 57 (13) years, disease duration of 12

Table 1 Patient characteristics by site, all self-reported

	All patients (n=274)	Bristol (n=33)	Vienna (n=51)	Amsterdam (n=54)	Paris (n=43)	Copenhagen (n=43)	USA (n=50)
Gender (% female)	75	73	75	76	79	61	82
Age in years (mean (SD))	57 (13)	62 (14)	54 (14)	60 (12)	53 (13)	56 (13)	56 (11)
Disease duration in years (mean, SD)	12 (9)	8 (8)	11 (9)	14 (12)	12 (8)	13 (9)	11 (7)
Experience with remission (% yes)	74	52	88	76	72	91	60
Currently in remission (% yes)	38	18	45	46	51	44	16
Self-reported deformities (% yes)	41	97	41	54	19	30	52
Pain (VAS 1–10)	3.5 (2.7)	5 (2.3)	2.4 (2.6)	3.1 (2.7)	3.4 (2.4)	2.6 (2.5)	5.0 (2.7)
PtGA (VAS 1–10)	3.6 (2.7)	4.9 (2.3)	3.2 (2.8)	3.1 (2.5)	3.2 (2.6)	2.3 (2.0)	4.9 (2.9)
RAPID3 (0–30)	8.9 (6.4)	13.6 (5.2)	6.4 (6.2)	8.3 (6.3)	8.0 (5.4)	6.2 (5.5)	12.6 (6.2)
RAPID3 near remission (% yes)	30	3	51	37	23	44	8
Ability to distinguish pain due to inflammation versus damage (% yes)	60	52	68	70	51	62	54

PtGA, patient global assessment; RAPID3, Routine Assessment of Patient Index Data 3; VAS, Visual Analogue Scale.

(9) years with 10% disease duration <2 years and self-reported erosive disease of 41%. Remission as reported by the patient was present in 38% and according to the RAPID3 in 30%. Concordance was reasonably good: of the patients in self-reported remission, 61% were in RAPID3 remission; of the patients in RAPID3 remission, 79% were in self-reported remission (observed agreement, 79%; κ 0.54).

In the subgroup of patients with available clinical data (n=119), 42% were both in self-reported and in RAPID3 remission, and 24% in ACR/EULAR remission. Here concordance was good only in one direction: of patients in ACR/EULAR remission 86% were in self-reported remission (observed agreement, 76%; κ 0.48) and 97% were in RAPID3 remission (observed agreement, 81%; κ 0.59); but only 50% of patients in self-reported remission and 56% of patients in RAPID3 remission were in ACR/EULAR remission (table 2). Of the patients with available CDAI (n=47), CDAI remission was present in 21%. Of these, all patients were in RAPID3 remission, 80% was in self-perceived remission (observed agreement, 70%; κ 0.35) and 70% in ACR/EULAR remission (observed agreement, 81%; κ 0.48).

Patients that considered themselves in remission (n=103) had a mean patient global assessment of disease activity of 1.5 (1.5) and disease duration of 10 (7) years) in contrast to 4.8(2.5) and 13(10) years for those not in self-perceived remission (n=171).

Most important domains

The most often-mentioned domains in the top 3 were as follows: pain (67%), fatigue (33%) and independence (19%) (table 3).

Only one domain, 'The way other people see me', was regarded by more than 30% of patients (59%) as 'not important' and was removed from further analyses. The percentage of patients that choose a certain domain in their top 3 is shown in online supplementary file II. Domains that were most frequently rated as 'essential' to characterise a period of remission were highly similar: pain (60%), being mobile (52%), physical function (51%), being independent (47%) and fatigue (41%) (table 4).

Pain needed to be less (13%), almost gone (42%) or gone (45%) to reflect remission. Similar patterns were seen for fatigue (23%, 40%, 37%). Independence needed to be better (16%), almost normal (31%) or normal (53%), with similar patterns for mobility (16%, 35%, 49%) and physical functioning (14%, 29%, 57%).

Stratifications

Age

Importance of pain and fatigue were similar in patients under and over 50 years of age; however, independence was reported more frequently in the top 3 by patients over 50 (24%) as compared with patients under 50 years of age (12%) (p=0.03).

Table 2 Concordance between two patient-derived remission definitions and the ACR/EULAR remission definition (count (percentage))

	Self-reported remission			ACR/EULAR remission		
	Yes	No	Total	Yes	No	Total
ACR/EULAR remission						
Yes	25 (51)	4 (6)	29 (24)			
No	24 (49)	66 (94)	90 (76)			
Total	49 (100)	70 (100)	119 (100)			
RAPID3 remission						
Yes	63 (61)	17 (10)	80 (30)	28 (97)	21 (24)	49 (42)
No	40 (39)	151 (90)	191 (70)	1 (3)	67 (76)	68 (58)
Total	103 (100)	168 (100)	271 (100)	29 (100)	88 (100)	117 (100)
Yes	8 (40)	2 (7)	10 (21)	7 (54)	3 (9)	10 (21)
CDAI*						
No	12 (60)	25 (93)	37 (79)	6 (46)	31 (91)	37 (79)
Total	20 (100)	27 (100)	47 (100)	13 (100)	34 (100)	47 (100)

*No data shown for CDAI versus RAPID3, as overlap was 100%.
ACR, American College of Rheumatology; CDAI, Clinical Disease Activity Index; EULAR, European League Against Rheumatism; RAPID3, Routine Assessment of Patient Index Data 3.

Disease duration

Importance of pain and fatigue were similar in patients with early (≤2 years) versus long-standing disease; numerically, independence was reported more frequently in the top 3 by patients with long-standing disease (21%) as compared with patients with early disease (12%) (p=0.40).

Country

There were slight differences in the choice of a top 3 in different countries (table 3), although pain was the number 1 domain in all countries. Pain was mentioned in the top 3 by 65%, 69%, 52%, 74%, 92% and 79% in the UK, Austria, the Netherlands, France, the USA and Denmark, respectively (χ^2 p=0.001). Likewise, fatigue was among the top 3 (48%, 16%, 30%, 51%, 40%, 29%, 35% (χ^2 p=0.005)) and to a lesser extent independence (39%, 20%, 27%, 21%, 6%, 17% and 21% (χ^2 p=0.01)). Bristol, Amsterdam and Paris had the same top 3 as that of the total group, but in Vienna, fatigue was replaced by stiffness and in both the USA and Copenhagen, independence was replaced by swelling.

Domains that were most frequently rated as 'essential' to characterise a period of remission were the same in the total group and in Bristol, while in Vienna fatigue was replaced by mental power, in Amsterdam fatigue was replaced by activities of daily living, in Paris fatigue was replaced by family role, in the USA physical function was replaced by mobility and in Copenhagen fatigue and independence were replaced by work and activities of daily living.

Version of the survey

The majority completed V.1 of the survey (79%). Patients that completed V.1 of the survey (appendix 1) reported pain in their top 3 more frequently (77%) compared with patients that completed V.2 of the survey (54%) (p=0.02). A similar pattern, although not significantly different, was seen for independence, reported by 19% in V.1 and 29% in V.2. However, of the patients that completed V.2 of the survey, the most often-mentioned domains in the top 3 were still pain (55%), fatigue (33%) and independence (29%).

Other

There were no differences in choice of a top 3 between male and female patients; between patients that reported joints with versus without strongly reduced mobility, deformities or joint replacement surgery; between patients that reported to have been diagnosed with other diseases; or between patients that completed the survey during their clinic visit, compared with patients that completed the survey electronically (at home).

Table 4 gives an overview of the numerical differences of domains rated as essential for all stratifications.

DISCUSSION

This survey study identified the three most important domains of patient perceived remission, based on preceding qualitative research on the patient perspective on remission in RA: the absence or reduction of pain and fatigue and the improvement or maintenance of independence.

Pain is the most predominant PRO assessed in rheumatic diseases, present in all core sets and frequently inquired after in clinical practice.[29] Fatigue has been acknowledged as an essential PRO in recent years and is recommended to be reported in RA clinical trials;[11] however, independence is not a common PRO in rheumatology. We are not the first to report on independence as an important domain for patients with RA; a recent Czech study[30] reported a significant difference between the healthy population and patients with RA in level of independence. In addition, qualitative work performed in Bristol generated the Rheumatoid Arthritis Patient Priorities for Pharmacological Interventions core set to complement the existing professional core sets.[31] This patient-driven study also identified independence as one of eight priority outcomes of pharmacological interventions in RA. In a Swedish qualitative study evaluating the patient perspective on benefits of RA treatment, independence was identified as a theme covering domains as management of daily activities, care for oneself and for one's family, being able to work and enjoy leisure time.[32]

In contrast, in the context of our remission work, independence emerged as a separate domain, grouped in the theme 'decreased impact of RA'. When reviewing our qualitative work, the domain independence seems mainly related to physical functioning, that is, 'the ability to do things you have to do and not have to ask others to do things for you'. However, influence of other domains and factors like ability to work, performing one's family role and comorbidities is highly likely and warrants further study. Similar overlap might exist between other domains while this could have been avoided by grouping of domains into larger themes. Indeed, if we would have grouped the domains socialise, family role, work and leisure into one domain 'participation', the ratings for this grouped domain would have exceeded that of independence in two out of six sites. However, we have chosen to present all 26 domains in our survey, using patient quotes from focus groups, so that patients

Table 3 Most important domains by site

	Bristol	Vienna	Amsterdam	Paris	Copenhagen	USA
Top 3						
1	Pain	Pain	Pain	Pain	Pain	Pain
2	Fatigue	Independence	Fatigue	Fatigue	Fatigue	Fatigue
3	Independence	Stiffness	Independence	Independence	Swelling	Swelling
Domains rated as essential (%)						
1	Pain (67)	Independence (53)	Pain (57)	Pain (72)	Pain (51)	Pain (70)
2	Fatigue (66)	Phys func (51)	Phys func (46)	Mobility (72)	Phys func (49)	Fatigue (60)
3	Mobility (66)	Pain (47)	ADL (43)	Phys func (67)	Work (49)	Independence (58)
4	Independence (50)	Mobility (43)	Independence (41)	Family (56)	Mobility (46)	Swelling (56)
5	Phys func (50)	Mental (43)	Mobility (39)	Indep (56)	ADL (44)	Mobility (54)

ADL, activities of daily living; Phys func., physical functioning.

Table 4 Percentage of patients that rate a domain as essential to be reduced or absent/improved to define remission

		Age		Gender		Disease duration		Comorbidities		Damage		Place of filling out		Version	
	Total	≤50	>50	Male	Female	≤2 years	>2 years	No	Yes	No	Yes	Clinic	Home	1	2
N	274	83	190	70	204	27	244	131	143	160	113	97	177	216	58
Pain	**60**	**71**	**56**	**55**	**62**	**57**	**61**	**62**	**59**	**59**	**62**	**56**	**62**	**62**	**54**
Fatigue	**41**	**49**	**38**	33	**44**	37	**42**	37	**45**	**43**	38	37	**44**	**41**	**42**
Physical funcion	**51**	**58**	**48**	**49**	**52**	**59**	**51**	**53**	**50**	**52**	**51**	**52**	**51**	**52**	**47**
Independence	**47**	**55**	**43**	**43**	**49**	**44**	**48**	**45**	**50**	**49**	**45**	**41**	**51**	**48**	**44**
Mobility	**52**	**60**	**48**	**43**	**55**	**44**	**53**	**53**	**51**	**57**	**46**	**55**	**51**	**55**	**42**
Swelling	38	**48**	34	30	**41**	**41**	39	39	38	**41**	35	37	39	38	**40**
Morning stiffness	24	29	22	20	25	0	15	24	24	21	28	22	25	26	18
Stiffness	30	36	27	27	31	15	31	31	30	28	33	30	30	31	28
Variability	23	26	21	20	23	15	23	21	24	23	23	21	23	23	19
Unpredictability	25	30	21	17	27	37	24	24	25	27	21	21	27	26	19
Flare	39	**51**	33	27	**42**	37	38	**41**	37	**42**	34	**43**	36	**41**	28
DMARD usage	31	30	32	29	32	26	31	29	33	33	30	38	27	34	20
Painkiller usage	26	31	24	21	28	26	26	27	25	26	27	29	24	28	19
Mood	21	21	21	9	25	30	20	17	24	23	18	18	22	20	23
Socialise	30	34	29	29	31	**50**	29	27	33	36	23	25	33	30	31
Strength	27	29	26	26	28	**44**	**53**	28	27	28	27	24	29	30	16
Grip	34	35	34	33	35	33	27	38	31	35	34	31	36	33	39
Fine motor skills	36	**43**	33	31	38	**48**	35	37	35	36	36	31	39	36	35
Activities of daily living	39	**46**	35	31	**41**	**48**	38	**42**	35	**42**	34	**41**	37	**43**	23
Family role	37	**54**	30	29	**40**	33	37	**41**	33	**45**	26	33	39	37	36
Work	33	**51**	25	30	34	27	33	33	32	**41**	22	33	33	35	25
Doing leisure time activities	27	36	23	27	27	30	27	30	25	33	20	25	29	30	18
Feeling as normal as other people around you	26	35	22	19	29	33	26	25	27	30	21	22	29	30	11
Being treated as normal as others by people around you	28	36	24	19	31	33	28	29	27	30	25	22	31	30	20
Mental power	**40**	**46**	37	29	**44**	**48**	39	**41**	39	**44**	35	36	**42**	**41**	35

DMARD, disease modifying antirheumatic drug.
Bold signifies frequencies that are at or above 40%.

could optimally relate to the domains. In the case of independence, this domain was chosen more often as a top 3 domain than related domains such as physical functioning, mobility, work or activities of daily living. However, these related domains scored high in the rating of essential domains, with mobility ranked second, physical function third, independence fourth and fatigue fifth. As our objective was to identify the three most important domains for patients with RA to define a state of remission, we have chosen to rely on the report of top 3 domains, knowing that the ranking of essential domains pointed in a highly similar direction and that independence covers aspects of function/mobility.

Some differences exist between different patient groups when stratified for age, gender, disease duration and the presence of comorbidities and damage; for example, performing one's family role was more important to younger patients compared with older patients. These differences were already predicted by patients during the qualitative phase and are confirmed by this survey study.

In addition, some differences exist in the procedure of filling out the survey; we sent out two versions of the survey with a different (random) order of the domains, to reduce any order effect on decision making. Although there was a difference in frequencies of domains reported in the top 3 between patients that completed V.1 and those that completed V.2 of the survey, the three most often reported domains were still pain, fatigue and independence. It seemed that the higher the domain was listed in the survey, the more likely that patients would include the domain in their top 3, even though the top 3 was a separate section at the end of the survey. However, as the top 3 remained unchanged across the different versions, this order effect had no influence on the patients' choice for the three most important domains.

It can be questioned whether other domains besides the 26 identified from focus groups in Bristol, Amsterdam and Vienna would have emerged when groups were organised in Paris, Copenhagen and the USA; the survey allowed patients to add additional domains. Seven patients used this opportunity, but this did not result in new domains, suggesting that the survey was comprehensive. Yet, this does not indicate generalisability of our results to countries beside those studied here.

The majority of patients indicated that they needed their symptoms to be 'gone' instead of 'less' or 'almost absent' to reflect remission. Yet, patients that indicated themselves as currently in remission had a mean patient global assessment of 1.5. It remains a point of discussion how this relates to the ACR/EULAR remission definition, which requires that the patient global assessment of disease activity can be no >1. Interpreting these data, one could argue that a score of <2 on a visual analogue scale from 0 to 10 equals 'gone' from a patient perspective. Alternatively, to adhere to the rule of 1, one could allow rounding to one for scores that go up to 1.5. Interesting is the contrast between patients' rating of fatigue as a top 3 priority, yet needing it to be less (23%), almost gone (40%) or gone (37%) to reflect remission. This is different for the other four important domains, in which the majority feels a certain domain needs to be gone, rather than almost gone or less. Perhaps fatigue is a domain that seriously impacts the lives of people with RA, but is somehow more manageable or acceptable than for example pain.

Strengths of this study include the geographical spread of the patient sample as well as the dual analysis of domains, either within the top 3 of importance or rated most frequently as essential. The robustness of the results across countries and response modes enhances the reliability of the results.

A weakness of this study is that the response rate was not registered at every site, which might indicate that a selection of highly motivated patients was studied. When looking at the demographic characteristics, a heterogeneous pattern is observed, with a wide range in patients' age, disease duration and remission experience. The size of our study population does not allow for extensive subgroup analysis. Yet the uniformity in a choice for a top 3 of patient important domains strengthens the importance of these three domains and suggests that these are—to a certain level—independent of other patient characteristics.

Another weakness of this study is the absence of extensive clinical data collection, which prohibited us from linking our results with other frequently used disease activity measures like the Clinical and Simplified Disease Activity Indices (except in a few patients). However, this work served primarily to identify the common top 3 of patient domains that need to be studied further in comparison with clinical measures.

With this in mind, a validation study has been initiated, studying measurement instruments for the three most important domains of patient perceived remission identified in this study in relation to clinical outcome. It is anticipated that the results of the validation study will inform whether any of these three domains add important information to the ACR/EULAR remission criteria.

In summary, this survey study identified the three most important domains of patient perceived remission, based on preceding qualitative research on the patient perspective on remission in RA. Follow-up research has been initiated to identify valid measurement instruments for these domains and quantify the contribution to the ACR/EULAR remission criteria.

Author affiliations
[1]Amsterdam Rheumatology & Immunology Center, Amsterdam, The Netherlands
[2]Medical University of Vienna, Vienna, Austria
[3]University of the West of England, Bristol, UK
[4]Patient Partner, Bristol, UK
[5]Quintiles and Duke University School of Medicine, Durham, North Carolina, USA
[6]Department of Rheumatology, Sorbonne Universités, UPMC Univ Paris 06; AP-HP, Pitié Salpêtrière Hospital, Paris, France
[7]Copenhagen Center for Arthritis Research and DANBIO, Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark
[8]University of Adelaide, Adelaide, South Australia, Australia
[9]Patient Partner, Amsterdam, The Netherlands
[10]University of Bristol, Bristol, UK
[11]University of Ottawa, Ottawa, Canada
[12]Department of Epidemiology & Biostatistics, VUmc, Amsterdam, The Netherlands

Twitter Follow Caroline Flurey at @C_Flurey

Acknowledgements The authors thank the OMERACT 'Patient perspective on remission in rheumatoid arthritis'—working group for their participation in this work.

Funding This work was funded by a grant from EULAR.

Competing interests None.

Ethics approval Medical Ethical Committees of all involved study sites: VU University Medical Center/Reade in Amsterdam, the Netherlands; Medical University of Vienna, Austria; Bristol Royal Infirmary in Bristol, United Kingdom (UK); Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark; Université Pierre et Marie Curie and Hopital Pitie-Salpetriere Paris in Paris.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Boers M, Tugwell P, Felson DT, *et al*. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol Suppl* 1994;41:86–9.
2 Van der Heijde DM, van 't Hof M, van Riel PL, *et al*. Development of a disease activity score based on judgment in clinical practice by rheumatologists. *J Rheumatol* 1993;20:579–81.

3 Pincus T, Amara I, Segurado OG, *et al*. Relative efficiencies of physician/assessor global estimates and patient questionnaire measures are similar to or greater than joint counts to distinguish adalimumab from control treatments in rheumatoid arthritis clinical trials. *J Rheumatol* 2008;35:201–5.
4 Sokka T, Häkkinen A, Krishnan E, *et al*. Similar prediction of mortality by the health assessment questionnaire in patients with rheumatoid arthritis and the general population. *Ann Rheum Dis* 2004;63:494–7.
5 Radner H, Alasti F, Smolen JS, *et al*. Physical function continues to improve when clinical remission is sustained in rheumatoid arthritis patients. *Arthritis Res Ther* 2015;17:203.
6 Aletaha D, Alasti F, Smolen JS. Chronicity of rheumatoid arthritis affects the responsiveness of physical function, but not of disease activity measures in rheumatoid arthritis clinical trials. *Ann Rheum Dis* 2015;74:532–7.
7 Hewlett S, Carr M, Ryan S, *et al*. Outcomes generated by patients with rheumatoid arthritis: how important are they? *Musculoskeletal Care* 2005;3:131–42.
8 Minnock P, Kirwan J, Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. *Rheumatology (Oxford)* 2009;48:1533–6.
9 Nicklin J, Cramp F, Kirwan J, *et al*. Collaboration with patients in the design of patient-reported outcome measures: capturing the experience of fatigue in rheumatoid arthritis. *Arthritis Care Res* 2010;62:1552–8.
10 Wells G, Li T, Maxwell L, *et al*. Responsiveness of patient reported outcomes including fatigue, sleep quality, activity limitation, and quality of life following treatment with abatacept for rheumatoid arthritis. *Ann Rheum Dis* 2008;67:260–5.
11 Kirwan JR, Newman S, Tugwell PS, *et al*. Progress on incorporating the patient perspective in outcome assessment in rheumatology and the emergence of life impact measures at OMERACT 9. *J Rheumatol* 2009;36:2071–6.
12 Aletaha D, Landewe R, Karonitsch T, *et al*. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Ann Rheum Dis* 2008;67:1360–4.
13 Gossec L, Dougados M, Rincheval N, *et al*. Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. *Ann Rheum Dis* 2009;68:1680–5.
14 Gossec L, de Wit M, Kiltz U, *et al*., EULAR PsAID Taskforce. A patient-derived and patient-reported outcome measure for assessing psoriatic arthritis: elaboration and preliminary validation of the Psoriatic Arthritis Impact of Disease (PsAID) questionnaire, a 13-country EULAR initiative. *Ann Rheum Dis* 2014;73:1012–19.
15 Sokka T, Hetland ML, Mäkinen H, *et al*. Remission and rheumatoid arthritis: data on patients receiving usual care in twenty-four countries. *Arthritis Rheum* 2008;58:2642–51.
16 van Tuyl LH, Lems WF, Voskuyl AE, *et al*. Tight control and intensified COBRA combination treatment in early rheumatoid arthritis: 90% remission in a pilot trial. *Ann Rheum Dis* 2008;67:1574–7.
17 Wells GA, Boers M, Shea B, *et al*. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005;32:2016–24.
18 Smolen JS, Aletaha D, Bijlsma JW, *et al*. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;69:631–7.
19 van Tuyl LH, Vlad SC, Felson DT, *et al*. Defining remission in rheumatoid arthritis: results of an initial American College of Rheumatology/European League Against Rheumatism consensus conference. *Arthritis Rheum* 2009;61:704–10.
20 Felson DT, Smolen J, Wells GA, *et al*. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011;63:573–86.
21 van Tuyl LH, Felson DT, Wells G, *et al*. Evidence for predictive validity of remission on long-term outcome in rheumatoid arthritis: a systematic review. *Arthritis Care Res (Hoboken)* 2010;62:108–20.
22 van Tuyl LH, Smolen JS, Wells GA, *et al*. Patient perspective on remission in rheumatoid arthritis. *J Rheumatol* 2011;38:1735–8.
23 van Tuyl LH, Hewlett S, Sadlonova M, *et al*. The patient perspective on remission in rheumatoid arthritis: 'You've got limits, but you're back to being you again'. *Ann Rheum Dis* 2015;74:1004–10.
24 Aletaha D, Neogi T, Silman AJ, *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;62:2569–81.
25 Goel N, Cummins G, Downing J, *et al*. Rapid acquisition of data on the patient perspective in rheumatoid arthritis through a digital portal. *Ann Rheum Dis* 2015;74(Suppl 2):557.
26 van Tuyl LH, Sadlonova M, Davis B, *et al*. Remission in rheumatoid arthritis: working toward incorporation of the patient perspective at OMERACT 12. *J Rheumatol* 2016;43:203–7.
27 Pincus T, Swearingen CJ, Bergman M, *et al*. RAPID3 (Routine Assessment of Patient Index Data 3), a rheumatoid arthritis index without formal joint counts for routine care: proposed severity categories compared to disease activity score and clinical disease activity index categories. *J Rheumatol* 2008;35:2136–47.
28 Aletaha D, Nell VP, Stamm T, *et al*. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005;7:R796–806.
29 van Tuyl LH, Boers M. Patient-reported outcomes in core domain sets for rheumatic diseases. *Nat Rev Rheumatol* 2015;11:705–12.
30 Tóthová V, Bártlová S, Dolák F, *et al*. Quality of life in patients with chronic diseases. *Neuro Endocrinol Lett* 2014;35(Suppl 1):11–18.
31 Sanderson T, Morris M, Calnan M, *et al*. Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. *Arthritis Care Res (Hoboken)* 2010;62:647–56.
32 Ahlmén M, Nordenskiöld U, Archenholtz B, *et al*. Rheumatology outcomes: the patient's perspective. A multicentre focus group interview study of Swedish rheumatoid arthritis patients. *Rheumatology (Oxford)* 2005;44:105–10.

OPEN ACCESS

EXTENDED REPORT

Cytosolic 5′-nucleotidase 1A autoantibody profile and clinical characteristics in inclusion body myositis

J B Lilleker,[1,2] A Rietveld,[3] S R Pye,[1] K Mariampillai,[4] O Benveniste,[4] M T J Peeters,[3] J A L Miller,[5] M G Hanna,[6] P M Machado,[6,7] M J Parton,[6] K R Gheorghe,[8] U A Badrising,[9] I E Lundberg,[8] S Sacconi,[10] M K Herbert,[11] N J McHugh,[12] B R F Lecky,[13] C Brierley,[14] D Hilton-Jones,[15] J A Lamb,[16] M E Roberts,[2] R G Cooper,[16,17,18] C G J Saris,[3] G J M Pruijn,[11] H Chinoy,[1,18,19] B G M van Engelen,[3] On behalf of all UKMYONET contributors

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210282).

For numbered affiliations see end of article.

Correspondence to
Dr J B Lilleker, Centre for Musculoskeletal Research, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, M13 9PT, UK; james.lilleker@manchester.ac.uk

JBL and AR are joint first authors.
HC and BGMvE are joint last authors.

Received 29 July 2016
Revised 7 October 2016
Accepted 5 November 2016
Published Online First: 25 January 2017

ABSTRACT

Objectives Autoantibodies directed against cytosolic 5′-nucleotidase 1A have been identified in many patients with inclusion body myositis. This retrospective study investigated the association between anticytosolic 5′-nucleotidase 1A antibody status and clinical, serological and histopathological features to explore the utility of this antibody to identify inclusion body myositis subgroups and to predict prognosis.

Materials and methods Data from various European inclusion body myositis registries were pooled. Anticytosolic 5′-nucleotidase 1A status was determined by an established ELISA technique. Cases were stratified according to antibody status and comparisons made. Survival and mobility aid requirement analyses were performed using Kaplan-Meier curves and Cox proportional hazards regression.

Results Data from 311 patients were available for analysis; 102 (33%) had anticytosolic 5′-nucleotidase 1A antibodies. Antibody-positive patients had a higher adjusted mortality risk (HR 1.89, 95% CI 1.11 to 3.21, p=0.019), lower frequency of proximal upper limb weakness at disease onset (8% vs 23%, adjusted OR 0.29, 95% CI 0.12 to 0.68, p=0.005) and an increased prevalence of excess of cytochrome oxidase deficient fibres on muscle biopsy analysis (87% vs 72%, adjusted OR 2.80, 95% CI 1.17 to 6.66, p=0.020), compared with antibody-negative patients.

Interpretation Differences were observed in clinical and histopathological features between anticytosolic 5′-nucleotidase 1A antibody positive and negative patients with inclusion body myositis, and antibody-positive patients had a higher adjusted mortality risk. Stratification of inclusion body myositis by anticytosolic 5′-nucleotidase 1A antibody status may be useful, potentially highlighting a distinct inclusion body myositis subtype with a more severe phenotype.

INTRODUCTION

Inclusion body myositis (IBM) is an acquired muscle disease that most commonly affects males aged over 45 years. Along with polymyositis (PM) and dermatomyositis (DM), IBM is usually classified as one of the idiopathic inflammatory myopathies. However, IBM differs in comparison with PM and DM, as sustained responses to immunosuppression are not seen, and histologically it is associated with significant degenerative features.[1–3] Clinically, IBM is characterised by asymmetric weakness, notably of finger flexors and knee extensors. Weakness in other muscle groups occurs frequently, including bulbar, facial and axial muscles.[4 5] The slowly progressive course leads to cumulative disability, although overall life expectancy is unaffected.[6–8]

The diagnosis of IBM relies upon a combination of clinical and laboratory findings as defined in various diagnostic criteria (eg, Medical Research Council (MRC), Griggs *et al* and the European Neuromuscular Centre (ENMC) criteria).[9–11] However, certain histopathological findings may only become detectable as the disease progresses, and therefore patients with early disease may not fulfil definite diagnostic criteria and can be excluded from clinical trials.[12] The average delay between disease onset and diagnosis is around 5 years, and IBM is frequently misdiagnosed initially as PM, resulting in the unnecessary use of potentially harmful treatments, such as high-dose glucocorticoids.[8 13–15]

In IBM, autoantibodies directed against cytosolic 5′-nucleotidase 1A (cN-1A) have recently been identified. It is suggested that these may support the diagnostic process, as well as potentially providing clues as to disease pathogenesis.[16 17] However, uncertainties regarding the usefulness of anti-cN-1A autoantibody testing in clinical practice remain. This is particularly true with regard to patient stratification and prognosis, where the few studies that have compared clinical and histopathological features of antibody-positive versus antibody-negative patients with IBM have produced conflicting results in some cases.[18 19] In order to explore further the usefulness of anti-cN-1A antibody testing to facilitate IBM subgroup classification, we conducted a retrospective Europe-wide study correlating clinical, serological and histopathological features in a large cohort of patients with IBM stratified by anti-cN-1A antibody status.

PATIENTS AND METHODS

Study cohort

Pooled IBM case data from four European countries were used. Researchers based in Nijmegen,

To cite: Lilleker JB, Rietveld A, Pye SR, *et al*. *Ann Rheum Dis* 2017;**76**:862–868.

The Netherlands, coordinated data collection from The Netherlands, France and Sweden. Data collection in the UK was coordinated by researchers based in Manchester, UK.

Study inclusion criteria

Included cases met either the MRC ('pathologically defined', 'clinically defined' or 'possible'), Griggs *et al* ('definite' or 'possible') or ENMC ('clinicopathologically defined', 'clinically defined' or 'probable') diagnostic criteria for IBM and had sera available for anti-cN-1A antibody testing.[9 11]

Data collection methodology

Swedish, French and Dutch ('non-UK') patients were identified from clinical databases. Researchers blinded to anti-cN-1A antibody status (AR, MTJP, KRG, KM) reviewed the medical records and retrospectively completed a standardised data collection pro forma. UK patients were identified from six centres contributing to the UKMYONET research study, coordinated by The University of Manchester. As part of this study, data are captured using a standardised pro forma at the time of study recruitment (ie, before serological test results are available).[20 21] Those recruiting patients are asked to record clinical features present at disease onset and features present at the time of recruitment. Some additional fields (to match data from the non-UK cohort) and missing data were collected retrospectively. Copies of pro forma used are contained in online supplementary appendix 1. The datasets were merged and cleaned by a researcher blinded to anti-cN-1A status (JBL).

Clinical data

Data collected included demographic, clinical (eg, distribution of weakness, presence of dysphagia, comorbidities), laboratory findings (creatine kinase (CK) levels, muscle biopsy features, serological testing), comorbidity, mobility aid usage and mortality. In most cases, data were available regarding features present at disease onset and at the time of last patient review (or recruitment to the UKMYONET study in the case of the UK cohort). In all cases, 'disease onset' refers to the initial date that symptoms of IBM were noted, as reported by the patient. 'Disease duration' is defined as the period between disease onset and the date of anti-cN-1A antibody testing. Regarding mortality, in the non-UK cohort, the primary cause of death was categorised by review of the patient's medical records as either 'respiratory', 'cardiac', 'cerebrovascular', 'malignancy' or 'other'. In the UK cohort, additional mortality and comorbidity statistics were obtained from the UK Health and Social Care Information Centre, including coded data regarding the cause of death where applicable. The cause of death in these cases was assessed and assigned to the same categories as the non-UK cohort.

Histopathology

For all cases, the histopathology biopsy report performed at initial diagnostic interrogation was reviewed, and the presence of certain specific features determined from the report text. The reporting histopathologists were blinded to the anti-cN-1A antibody status of each patient at the time of reporting. Cytochrome oxidase (COX) deficient fibres in the biopsy sample were recorded as 'excessive' if the reporting histopathologist indicated that numbers were adjudged higher than expected, according to the patient's age. In some cases, the date that the biopsy was performed was not available. In such instances, this was assumed to be the same as the date of diagnosis.

cN-1A analysis

All sera were analysed at the Department of Biomolecular Chemistry in Nijmegen by ELISA, with the three synthetic peptides containing cN-1A autoepitopes previously identified by overlapping peptide microarray analyses.[16] Signals were quantified by determining optical densities at 450 nm (OD450) using methods previously described and defined as seropositive if the OD450 value was greater than or equal to the established cut-off value for the corresponding peptide.[22]

Other serological testing

Data regarding the presence of myositis-specific antibodies (MSAs) and myositis-associated antibodies (MAAs) were collected where available. For the non-UK patients, data were obtained from results available in the medical records, and methodology of testing was unique to each centre. MSAs and MAAs in the whole UK cohort were screened by immunoprecipitation at the University of Bath (Bath, UK) using previously described standardised methodology.[23] 'Weak positive' results were assumed to be negative for the purpose of this study.

Statistical analysis

The per-subject sum of all recorded comorbidities (of autoimmune disease, cardiovascular disease (including hypertension) and malignancy) was calculated. Current or previous smoking was also treated as a comorbidity for the purposes of this analysis. According to the number of these factors present, each patient was then assigned a comorbidity score of 0, 1 or 2 or more for use in regression. Differences in demographic features, comorbidities, clinical features, autoantibody status and muscle biopsy features between anti-cN-1A antibody positive and negative patients were assessed using logistic regression. In order to test the effect of potential confounders, adjusted (multivariable) logistic regression models were produced when unadjusted analysis had suggested a significant difference (defined as $p<0.05$).

The impact of anti-cN-1A antibody status on survival and mobility aid requirement was assessed using Kaplan-Meier curves, log-rank testing and Cox proportional hazards regression modelling. In both cases, the start of the surveillance period was the date of disease onset. For the mobility aid analysis, subjects exited the model at the time of mobility aid requirement or at the time they were last known to have not required one. For the survival analysis, subjects exited the model at the time of death or at the time they were last known to have been alive. Each Cox regression model included adjustment for age of disease onset, gender and comorbidities. Other variables were added to the models if there was an a priori assumption that a relationship between anti-cN-1A antibody status and the outcome variable was likely to exist. For example, a higher incidence of anti-cN-1A antibodies in those with Sjögren's syndrome is reported, a more prominent bulbar involvement in anti-cN-1A positive patients with IBM has been described and a correlation between COX deficiency and more advanced age at biopsy could exist.[18 22 24] Therefore, models with additional adjustment for such variables were created.

The analysis plan specifically omitted correction for multiple testing due to the highly conservative nature of such methods which would risk elimination of potentially useful information which was sought to be retained, given the exploratory nature of this study. Data were processed and analysed using Stata for Windows V.13.0 (College Station, Texas, USA). Kaplan-Meier curves were generated using GraphPad Prism V.6 (GraphPad Software).

RESULTS

After screening databases in the four involved countries, 311 patients meeting the study inclusion criteria were selected for further analysis (45% from the UK, 55% non-UK). Overall, 33% (102/311) were positive for the anti-cN-1A antibody. Table 1 shows the IBM diagnostic criteria met according to anti-cN-1A antibody status. No relationship between a diagnostic classification of 'possible' IBM versus 'definite' (for Griggs *et al* criteria) or 'pathologically/clinically defined' (for MRC criteria) IBM and anti-cN-1A antibody status was found (for MRC criteria, OR 0.85, 95% CI 0.48 to 1.49, p=0.565; for Griggs *et al* criteria, OR 0.70, 95% CI 0.36 to 1.36, p=0.292; analysis not performed for ENMC criteria as all anti-cN-1A antibody positive patients met the definition of 'definite' IBM). No difference was found in the interval between disease onset and the time of antibody testing between seropositive and seronegative groups (8.29 years (IQR 4.96–11.95) in the seropositive group vs 7.57 years (IQR 4.94–11.18) in the seronegative group, OR 1.01, 95% CI 0.97 to 1.06, p=0.604).

Demographics and comorbidities

No statistically significant differences were identified in demographic characteristics (including gender, age at disease onset and age at diagnosis), CK levels, smoking history or comorbidities between the anti-cN-1A antibody positive and negative groups (table 2). Non-significant trends were observed in age at disease onset and age at diagnosis (which appeared lower in the antibody-negative group) or the presence of other autoimmune diseases (which appeared more common in the antibody-positive group).

Survival

Of the whole cohort of 311 patients, 70 deaths were recorded (31/102 (30%) in the anti-cN-1A antibody positive group and 39/209 (19%) in the negative group). The mean age of death overall was 77.8 years (SD=8.2), with no significant difference detected according to anti-cN-1A antibody status (77.0 years (SD=7.7) in the seropositive group vs 78.4 years (SD=8.6) in the seronegative group, OR 0.98, 95% CI 0.92 to 1.04, p=0.482). The cause of death was known in 63% (44 of 70) of cases. An excess of deaths as a result of respiratory disease was evident in the anti-cN-1A antibody positive group (16/25 (64%) in the anti-cN-1A antibody positive group and 9/25 (36%) in the negative group, OR 4.23, 95% CI 1.79 to 9.97, p=0.001). Adjusted analysis was not performed here due to the low numbers available for analysis. Death from other causes (cardiac, cerebrovascular, malignancy and other causes) did not differ between anti-cN-1A antibody positive and negative groups.

Data from 300 patients, where the date of disease onset and date of last follow-up (or date of death) were known, were available for further analysis. This included 66 of those that had died (66/70, 94%) and comprised a total of 3550 patient-years of follow-up. The median survival in the anti-cN-1A antibody positive group was 17.6 years compared with 24.2 years in the antibody-negative group, and the Kaplan-Meier curves were significantly different (log-rank p=0.045, figure 1).

In unadjusted analysis, compared with the antibody-negative group, anti-cN-1A antibody positive patients had a 65% increased risk of death (HR 1.65, 95% CI 1.01 to 2.70, p=0.047). After adjustment for age at disease onset, gender and comorbidities, the HR was 1.95 (95% CI 1.17 to 3.27, p=0.011). Furthermore, adding the presence of dysphagia to the model confirmed an independent association (HR 1.89, 95% CI 1.11 to 3.21, p=0.019).

Mobility

Data from 188 patients were available for this analysis. A total of 130 instances of mobility aid uptake were recorded, 81% (52/64) in the anti-cN-1A seropositive group and 63% (78/124) in the seronegative group. The overall median time between disease onset and use of a mobility aid was 8.0 years (IQR 4.6–11.0), with no significant difference between seropositive and seronegative groups (8.0 years (IQR 4.8–10.9), and 6.9 years (IQR 4.4–11.7), respectively; OR 1.01, 95% CI 0.94 to 1.08, p=0.883). Kaplan-Meier curves were not significantly different (log-rank p=0.090), so not shown. In unadjusted analysis, the HR for mobility aid requirement in the antibody-positive group was 1.35 (95% CI 0.95 to 1.93, p=0.097). After adjustment for age at disease onset, gender and comorbidities, the HR for mobility aid requirement was just outside the significance threshold (HR 1.42, 95% CI 0.99 to 2.04, p=0.056).

Clinical features

Table 3 demonstrates the clinical characteristics at disease onset and at last clinical review, stratified by anti-cN-1A antibody status. A significant association between the presence of proximal upper limb weakness at disease onset (not a typical feature of IBM) and being anti-cN-1A antibody negative was identified (OR 0.30 95% CI 0.13 to 0.71, p=0.006). This remained significant after adjustment for age at onset, gender and comorbidities (OR 0.29, 95% CI 0.12 to 0.68, p=0.005), thus potentially defining a more classical and homogenous IBM cohort in the anti-cN-1A antibody positive group. Data regarding the presence of facial weakness were less complete (n=90). Despite this, a significantly increased incidence of facial weakness was identified in the anti-cN-1A antibody positive group at last review (OR 2.60, 95% CI 1.07 to 6.29, p=0.034), which persisted after adjustment for age at onset, gender and comorbidities (OR 3.03, 95% CI 1.20 to 7.67, p=0.019).

Autoantibody associations

A significant association between seropositivity for anti-SSB (La) antibodies and anti-cN-1A antibodies was identified (OR 3.28, 95% CI 1.33 to 8.07, p=0.010) (table 4). However, adjusted

Table 1 Summary of diagnostic criteria met in patients included for analysis

Diagnostic criteria met	Anti-cN-1A positive (%)	Total (all patients)
Medical Research Council Criteria 2010[10]		
Pathologically defined IBM	13 (31.7)	41
Clinically defined IBM	39 (39.4)	99
Possible IBM	28 (33.3)	84
Griggs *et al*[9] Criteria		
Definite IBM	19 (40.4)	47
Possible IBM	61 (32.3)	189
European Neuromuscular Centre Criteria 1997[11]		
Definite IBM	7 (31.8)	22
Probable IBM	0 (0.0)	2
Total unique patients*	102 (32.8)	311

*Some patients fulfilled multiple diagnostic criteria. Not all patients were assessed by each criterion. Of the total, 152 patients met only one criterion, 143 patients met two criteria and 16 patients met all three criteria.
Anti-cN-1A, anticytosolic 5′-nucleotidase 1A; IBM, inclusion body myositis.

Table 2 Summary of demographic features, CK levels and comorbidities stratified by anti-cN-1A antibody status

	Anti-cN-1A positive	Anti-cN-1A negative	OR (95% CI)	p Value
Gender (n=311)				
Female (%)	42/102 (41.2)	84/209 (40.2)	Referent	–
Male (%)	60/102 (58.8)	125/209 (59.8)	0.96 (0.59 to 1.55)	0.868
Ethnicity (n=307)				
White (%)	97/101 (96.0)	199/206 (96.6)	Referent	–
Black (%)	2/101 (2.0)	4/206 (1.9)	1.03 (0.19 to 5.70)	0.977
Asian (%)	2/101 (2.0)	3/206 (1.5)	1.37 (0.23 to 8.32)	0.734
Other features				
Mean age in years at disease onset (SD) (n=301)	61.6 (9.7)	59.8 (9.5)	1.02 (0.99 to 1.05)	0.130
Mean age in years at diagnosis (SD) (n=305)	67.2 (9.3)	65.3 (9.5)	1.02 (1.00 to 1.05)	0.089
Disease duration in years at antibody testing (n=301)	Median 8.3 (IQR 5.0–12.0) Mean 9.0 (SD 5.5)	Median 7.6 (IQR 4.9–11.2) Mean 8.6 (SD 5.2)	1.01 (0.97 to 1.06)	0.604
Highest CK level recorded (n=223)	Median 629.0 (IQR 392–850) Mean 774.8 (SD 563.4)	Median 600.0 (IQR 400–1012) Mean 1097.2 (SD 2583.4)	1.00 (1.00 to 1.00)	0.318
Current or previous smoker (%) (n=189)	21/52 (40.4)	55/137 (40.2)	1.01 (0.53 to 1.94)	0.976
Comorbidities				
Autoimmune disease (including Sjögren's syndrome) (%) (n=244)	38/85 (44.7)	54/159 (34.0)	1.57 (0.92 to 2.70)	0.100
Of which, Sjögren's syndrome (%) (n=81)	6/33 (18.2)	8/48 (16.7)	1.11 (0.35 to 3.57)	0.859
Malignancy (%) (n=275)	12/85 (14.1)	33/190 (17.4)	0.78 (0.38 to 1.60)	0.501
Cardiovascular disease (%) (n=284)	31/91 (34.1)	64/193 (33.2)	1.04 (0.62 to 1.76)	0.880
Hypertension (%) (n=181)	29/60 (48.3)	54/121 (44.6)	1.16 (0.62 to 2.16)	0.638

'Disease duration in years at antibody testing' refers to the time period between disease onset and the date of anti-cN-1A antibody testing. n represents data available for analysis per variable (of a total of 311). p Value is derived from logistic regression.
Anti-cN-1A, anticytosolic 5′-nucleotidase 1A; CK, creatine kinase.

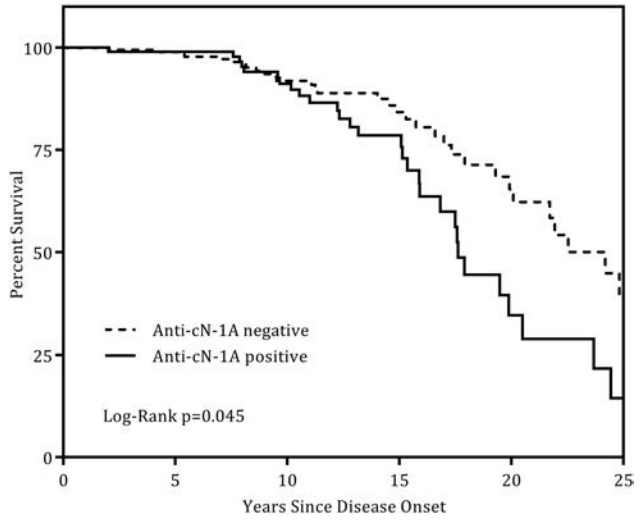


Figure 1 Kaplan-Meier survival curves stratified by anti-cN-1A antibody status. X-axis truncated at 25 years from disease onset.

analysis (for anti-SSA antibodies, presence of autoimmune disorders, age at onset, gender and comorbidities) did not confirm that this association was independent (OR 2.12, 95% CI 0.52 to 8.67, p=0.297).

Biopsy features

We identified a significant association between an excess of COX-deficient fibres on muscle biopsy and the presence of anti-cN-1A antibodies (OR 2.61, 95% CI 1.13 to 6.03, p=0.025) (table 5). In adjusted analysis (for age at disease onset, gender, comorbidities and age at biopsy), a significant independent association was confirmed (OR 2.80, 95% CI 1.17 to 6.66, p=0.020).

DISCUSSION

This multinational exploratory study represents the first of its kind to combine analysis of clinical, histopathological, other serological and mortality data in a large cohort of patients with IBM stratified according to anti-cN-1A antibody status. Our results will guide future confirmatory studies and highlight potential disease mechanisms warranting further evaluation. We found that the anti-cN-1A antibody positive group had a significantly increased mortality risk independent of age, gender, comorbidities and the presence of dysphagia. We also found a smaller proportion with proximal upper limb weakness at disease onset and an excess of COX-deficient fibres on muscle biopsy in the anti-cN-1A antibody positive group. An increased likelihood of having facial weakness and an association between antibody positivity and death from a respiratory cause was also observed, although the numbers assessed here were small. As in other studies, we did not find a relationship between disease duration and the likelihood of identifying anti-cN-1A antibodies.[18 19]

There are limited reports in the literature comparing the characteristics of patients with IBM with and without anti-cN-1A antibodies, amounting to 258 patients in four separate studies.[18 19 24 25] A small proportion of the cases analysed here was included in a previous analysis which did not focus on differences on clinical characteristics according to serotype.[22] Some authors identified no significant differences in the characteristics between cohorts, whereas others have suggested that the anti-cN-1A antibody positive group exhibits a more severe phenotype.[18 19] Lloyd *et al*[24] identified a lower incidence of rimmed vacuoles on biopsy in those without anti-cN-1A

Table 3 Clinical characteristics at disease onset and at last clinical review stratified by anti-cN-1A antibody status

Clinical feature	Anti-cN-1A positive (%)	Anti-cN-1A negative (%)	OR (95% CI)	p Value
At disease onset				
Proximal upper limb weakness (n=252)	7/84 (8.3)	39/168 (23.2)	0.30 (0.13 to 0.71)	0.006*
Proximal lower limb weakness (n=253)	65/85 (76.5)	122/168 (72.6)	1.23 (0.67 to 2.24)	0.510
Distal upper limb weakness (n=251)	22/83 (26.5)	40/168 (23.8)	1.15 (0.63 to 2.11)	0.641
Distal lower limb weakness (n=250)	7/83 (8.4)	20/167 (12.0)	0.68 (0.27 to 1.67)	0.398
Dysphagia (n=119)	15/36 (41.7)	23/83 (27.7)	1.86 (0.82 to 4.22)	0.136
Axial involvement (n=102)	0/30 (0.0)	3/72 (4.2)	1	–
Symmetrical weakness (n=97)	25/37 (67.6)	32/60 (53.3)	1.82 (0.78 to 4.29)	0.169
At last review				
Proximal lower limb weakness (n=137)	35/40 (87.5)	80/97 (82.5)	1.49 (0.51 to 4.35)	0.468
Distal upper limb weakness (n=135)	40/41 (97.6)	89/94 (94.7)	2.25 (0.25 to 19.86)	0.466
Distal lower limb weakness (n=125)	23/43 (53.5)	36/82 (43.9)	1.47 (0.70 to 3.08)	0.309
Dysphagia (n=303)	63/100 (63.0)	113/203 (55.7)	1.36 (0.83 to 2.22)	0.224
Facial weakness (n=90)	18/33 (54.6)	18/57 (31.6)	2.60 (1.07 to 6.29)	0.034†
Axial involvement (n=84)	9/26 (34.6)	10/58 (17.2)	2.54 (0.88 to 7.31)	0.084
Clinical evidence of polyneuropathy (n=103)	13/38 (34.2)	31/65 (47.7)	0.57 (0.25 to 1.31)	0.184

Figures in brackets represent *within antibody group* percentages. n represents data available for analysis per variable (of a total of 311). p Value is derived from logistic regression. Data regarding certain variables (proximal upper limb weakness, facial weakness, symmetrical weakness and clinical evidence of polyneuropathy) were only available at either disease onset or at last review.
*Adjusted (for age at disease onset, gender and comorbidities) OR 0.29, 95% CI 0.12 to 0.68, p=0.005.
†Adjusted (for age at disease onset, gender and comorbidities) OR 3.03, 95% CI 1.20 to 7.67, p=0.019.
Anti-cN-1A, anticytosolic 5′-nucleotidase 1A.

Table 4 Autoantibody profile stratified by anti-cN-1A antibody status

Antibody	Anti-cN-1A positive (%)	Anti-cN-1A negative (%)	OR (95% CI)	p Value
Antinuclear antibodies (n=132)	1/47 (2.1)	1/85 (1.2)	1.83 (0.11 to 29.88)	0.673
Anti-DNA antibodies (n=119)	3/42 (7.1)	1/77 (1.3)	5.85 (0.59 to 58.07)	0.132
Anti-Sm antibodies (n=97)	0/33 (0.0)	1/64 (1.6)	1	–
Antineutrophil cytoplasmic antibodies (n=96)	0/32 (0.0)	0/64 (0.0)	–	–
Antimitochondrial antibodies (n=128)	0/41 (0.0)	0/87 (0.0)	–	–
Antiextractable nuclear antigens antibodies (n=102)	4/34 (11.8)	5/68 (7.4)	1.68 (0.42 to 6.71)	0.463
Anti-SSA (Ro) (n=228)	19/76 (25.0)	22/152 (14.5)	1.97 (0.99 to 3.92)	0.054
Anti-SSB (La) (n=228)	13/76 (17.1)	9/152 (5.9)	3.28 (1.33 to 8.07)	0.010*
(U1)RNP antibodies (n=223)	1/74 (1.4)	0/149 (0.0)	1	–
Antitopoisomerase I (Scl70) (n=222)	0/72 (0.0)	0/150 (0.0)	–	–
Anti-Jo1 (n=228)	1/76 (1.3)	0/152 (0.0)	1	–
Other myositis-*specific* antibody (OMSA)† (n=193)	0/60 (0.0)	1/133 (0.8)	1	–
Other myositis-*associated* antibody (OMAA) (n=128)	0/41 (0.0)	0/87 (0.0)	–	–

Figures in brackets represent *within antibody group* percentages. n represents data available for analysis per variable (of a total of 311). p Value is derived from logistic regression.
*Adjusted (for anti-SSA antibodies, presence of autoimmune disorders, age at disease onset, gender and comorbidities) OR 2.11, 95% CI 0.52 to 8.67, p=0.297.
†One patient found positive for anti-SRP antibodies. In this case, no relevant clinical correlation was identified, and the relevance of this finding is uncertain.
Anti-cN-1A, anticytosolic 5′-nucleotidase 1A; OMAA, anti-Ku, anti-RNA polymerase I/II/III, anti-PM/SCL, anti-NOR90; OMSA, anti-TIF1 complex, anti-SAE, anti-NXP2, anti-MDA5, anti-SRP, anti-Mi-2, anti-PL12, anti-PL7, anti-EJ, anti-KS, anti-OJ, anti-Zo.

reactivity but with no clinical differences between the studied cohorts, findings that were not replicated here. A very recent study found no differences between 24 cN-1A seropositive and 45 seronegative patients with IBM regarding class II human leukocyte antigen (HLA) alleles and the presence of other antibodies.[25]

The simultaneous discovery of anti-cN-1A antibodies in 2011 by two independent research groups offers potential insights into the pathogenesis of IBM, and will contribute to the debate about the relative influence of the immune system and degeneration.[16 19 23] The presence of anti-cN-1A in other autoimmune diseases such as Sjögren's syndrome is also of interest as it might highlight shared underlying immune mechanisms across these diseases.[22] As with most other MSAs, further research is required to establish the mechanisms involved in anti-cN-1A reactivity in IBM.

Anti-cN-1A antibodies are present in the sera of 29%–52% of patients with IBM (33% in our cohort).[16 17] Higher proportions of anti-cN-1A antibody seropositivity in other studies (up to 72%) might be explained by different techniques used in different centres, by different cut-off levels for positivity or by differences in patient selection.[18] The current study used very strict cut-off values in ELISA testing.[26] In a recent study, anti-cN-1A antibodies were found in 37% of patients with IBM, compared with <5% in PM, DM and other neuromuscular disorders, highlighting a potential utility of using anti-cN-1A antibody

Table 5 Summary of muscle biopsy features stratified by anti-cN1-A antibody status

Biopsy feature	Anti-cN-1A positive (%)	Anti-cN-1A negative (%)	OR (95% CI)	p Value
Excess COX-deficient fibres (n=185)	53/61 (86.9)	89/124 (71.8)	2.61 (1.13 to 6.03)	0.025*
Ragged red fibres (n=164)	30/55 (54.6)	54/109 (49.5)	1.22 (0.64 to 2.34)	0.545
Atrophic fibres (n=176)	59/69 (85.5)	98/107 (91.6)	0.54 (0.21 to 1.41)	0.209
Inflammation (n=290)	94/96 (97.9)	193/194 (99.5)	0.24 (0.02 to 2.72)	0.251
MHC I upregulation (n=198)	67/69 (97.1)	124/129 (96.1)	1.35 (0.26 to 7.15)	0.724
Necrosis (n=136)	40/50 (80.0)	61/86 (70.9)	1.64 (0.71 to 3.78)	0.246
Mononuclear infiltrate (n=224)	72/74 (97.3)	143/150 (95.3)	1.76 (0.36 to 8.70)	0.487
Invasion of non-necrotic fibres ('partial invasion') (n=95)	21/30 (70.0)	48/65 (73.9)	0.83 (0.32 to 2.15)	0.696
Rimmed vacuoles (n=257)	77/88 (87.5)	143/169 (84.6)	1.27 (0.60 to 2.72)	0.533
Protein deposits† (n=128)	24/44 (54.6)	53/84 (63.1)	0.70 (0.34 to 1.47)	0.349
Microfilaments‡ (n=81)	9/24 (37.5)	24/57 (42.1)	0.83 (0.31 to 2.20)	0.700

Figures in brackets represent *within antibody group* percentages. n represents data available for analysis per variable (of a total of 311). p Value is derived from logistic regression.
*Adjusted (for age at disease onset, gender and comorbidities) OR 2.60, 95% CI 1.11 to 6.12, p=0.028. Adjusted (additionally for age at biopsy) OR 2.80, 95% CI 1.17 to 6.66, p=0.020.
†Includes amyloid (Congo Red or immunofluorescence), p62 (immunofluorescence) and TDP-43 (immunofluorescence).
‡15–21 nm tubulofilaments identified by electron microscopy.
Anti-cN-1A, anticytosolic 5′-nucleotidase 1A; COX, cytochrome oxidase; MHC, major histocompatibility complex.

testing to differentiate IBM and mimicking diagnoses.[22] However, the specificity of testing is limited by a high reactivity in some other autoimmune and connective tissue diseases (in 36% of patients with Sjögren's syndrome and in 20% with systemic lupus erythematosus).[22 24]

The higher frequency of COX-negative fibres, a feature of mitochondrial dysfunction, indicates possible differences in molecular pathways within the subgroups defined by anti-cN-1A antibody status. The reasons for increased mortality and the suggestion of increased risk of death from respiratory cause are unexplained, but these findings appear to agree with those of Goyal *et al*[18] who also found a more severe respiratory phenotype in the antibody-positive group. The lower frequency of proximal upper limb weakness at presentation in the anti-cN-1A antibody positive compared with antibody-negative group remains unexplained.

This study represents the largest cohort of patients with IBM and has only been achieved by an international collaborative effort. Established IBM diagnostic criteria were used to include patients for the analysis, a predefined set of clinical data was retrieved in each patient and all anti-cN-1A testing was performed in one laboratory. However, there remain a number of limitations. The study was retrospective and relied on the identification and recording of clinical characteristics by the treating physicians. In the UK cohort, the recruiting physician (the patient's treating consultant neurologist or rheumatologist) was asked to recall the symptoms that were present at the time of disease onset when completing the pro forma at the time of recruitment, and as such these details may be subject to recall bias. While efforts to minimise missing data were made, data were not complete for all study parameters in all cases, although there was no evidence to suggest that this occurred in a systematic way. Analysis involved pooling of data from different cohorts. There is potential for differences in data collection methodology between cohorts (see online supplementary appendix 1) to reduce the reliability of our findings. However, a comparison of features between UK and non-UK cohorts where pooled data were analysed has revealed largely comparable findings (see online supplementary table S1). Overall, we feel that our pooled analysis has increased statistical power and reduced the likelihood of statistical errors occurring. Objective measurements of muscle strength (eg, dynamometry of the finger flexors) could have improved sensitivity of detection of weakness, but such methods were not available. Also, this study did not perform a reanalysis of muscle biopsy tissue. The cause of death was difficult to establish in some patients in the non-UK cohort, due to missing information in the medical records, and in the UK cohort due to an inability to match some patients to the nationally stored mortality data held by the UK Health and Social Care Information Centre.

In the future, anti-cN-1A autoantibody testing and anti-cN-1A autoantibody status could be used in the diagnostic workup of potential IBM cases, and there remains the opportunity to use anti-cN-1A antibody status in the construction of future diagnostic criteria for IBM. However, the results of the current study also suggest that distinct IBM subtypes may be identified according to anti-cN1-A antibody status. Therefore, serum anti-cN-1A testing might also be of use in the stratification of patients with IBM (eg, for clinical trials), rather than purely as a diagnostic biomarker. A large prospective study with a sufficient duration of follow-up might offer potential to further investigate the overall utility of anti-cN-1A antibody testing in the clinical and research settings.

CONCLUSION

In this exploratory study, comparison of patients with IBM with and without anti-cN-1A autoantibody reactivity identified differences in their mortality risk, clinical characteristics and histopathological findings. The largest study of its kind has demonstrated that anti-cN-1A antibody testing may, and over and above its diagnostic value, be clinically useful to define distinct IBM subtypes.

Author affiliations
[1]Centre for Musculoskeletal Research, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
[2]Greater Manchester Neurosciences Centre, Salford Royal NHS Foundation Trust, Stott Lane, Salford, UK
[3]Department of Neurology, Center for Neuroscience Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands
[4]Department of Internal Medicine and Clinical Immunology, La Pitié-Salpêtrière Hospital, AP-HP, INSERM U974, UPMC, Paris, France
[5]Department of Neurology, Royal Victoria Hospitals, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle, UK

[6]MRC Centre for Neuromuscular Diseases, Institute of Neurology, University College London, London, UK
[7]Centre for Rheumatology Research, University College London, London, UK
[8]Unit of Rheumatology, Department of Medicine, Karolinska University Hospital, Solna, Karolinska Institutet, Stockholm, Sweden
[9]Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands
[10]Peripheral Nervous System, Muscle and ALS Department, Université Côté Azure (UCA), Nice University Hospital, Nice, France
[11]Department of Biomolecular Chemistry, Radboud Institute for Molecular Life Sciences and Institute for Molecules and Materials, Radboud University, Nijmegen, The Netherlands
[12]Royal National Hospital for Rheumatic Diseases and Department of Pharmacy and Pharmacology, University of Bath, Bath, UK
[13]The Walton Centre NHS Foundation Trust, Fazakerley, Liverpool, UK
[14]Department of Neurology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
[15]Nuffield Department of Clinical Neurosciences, Oxford University Hospitals, Oxford, UK
[16]Centre for Integrated Genomic Medical Research, University of Manchester, Manchester, UK
[17]MRC-ARUK Institute for Ageing and Chronic Disease, University of Liverpool, Liverpool, UK
[18]Rheumatology Department, Salford Royal NHS Foundation Trust, Salford, UK
[19]NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK

Twitter Follow Hector Chinoy at @drhectorchinoy

Acknowledgements The authors thank Hazel Platt (Centre for Integrated Genomic Medical Research, University of Manchester) and Paul New (Salford Royal NHS Foundation Trust). The authors thank all of the patients and their families who contributed to this study.

Contributors Initiation and design of this research: *Non-UK*—CGJS, BGMvE and GJMP; *UK*—RGC, HC and JAL. Clinical data collection and processing: AR, JBL, MTJP, KM and KRG. Facilitation of clinical data collection, establishment of the cohorts, contribution of cases: UAB, OB, IEL, SS, HC, RGC, JALM, MGH, PMM, MJP, BRFL, CB, DH-J and MER. Establishment of the antibody detection method and laboratory analysis: MKH, BGMvE and GJMP. Statistical analysis: JBL and SRP. Draft manuscript preparation: AR and JBL. All authors were involved with the review of the manuscript and approved the final version.

Funding This study was supported in part by: the Prinses Beatrix Spierfonds (W.OR 12–15); Myositis UK; Arthritis Research UK (18474); Association Française Contre Les Myopathies; The European Union Sixth Framework Programme (project AutoCure; LSH-018661); European Science Foundation in the framework of the Research Networking Programme European Myositis Network; The Swedish Research Council. PMM was supported by a National Institute for Health Research (NIHR) Rare Diseases Translational Research Collaboration Fellowship. This report includes independent research supported by the NIHR Biomedical Research Unit cFunding Scheme. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. The UKMYONET project is supported by the Manchester Academic Health Sciences Centre (MAHSC).

Competing interests GJMP and BGMvE are inventors of a patent (EP20120740236) licensed to Euroimmun, and GJMP receives financial support from Euroimmun for his research programme. Leiden University Medical Center receives financial compensation from Novartis for the BYM338 clinical trials in IBM in which UAB is the principal investigator.

Ethics approval Local ethics committee of each of the participating centres.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Badrising UA, Maat-Schieman M, van Duinen SG, *et al*. Epidemiology of inclusion body myositis in the Netherlands: a nationwide study. *Neurology* 2000;55:1385–7.
2 Wilson FC, Ytterberg SR, St Sauver JL, *et al*. Epidemiology of sporadic inclusion body myositis and polymyositis in Olmsted County, Minnesota. *J Rheumatol* 2008;35:445–7.
3 Benveniste O, Stenzel W, Hilton-Jones D, *et al*. Amyloid deposits and inflammatory infiltrates in sporadic inclusion body myositis: the inflammatory egg comes before the degenerative chicken. *Acta neuropathol* 2015;129:611–24.
4 Mastaglia FL, Needham M. Inclusion body myositis: a review of clinical and genetic aspects, diagnostic criteria and therapeutic approaches. *J Clin Neurosci* 2015;22:6–13.
5 Oh TH, Brumfield KA, Hoskin TL, *et al*. Dysphagia in inflammatory myopathy: clinical characteristics, treatment strategies, and outcome in 62 patients. *Mayo Clin Proc* 2007;82:441–7.
6 Cox FM, Titulaer MJ, Sont JK, *et al*. A 12-year follow-up in sporadic inclusion body myositis: an end stage with major disabilities. *Brain* 2011;134(Pt 11):3167–75.
7 Rose MR, McDermott MP, Thornton CA, *et al*. A prospective natural history study of inclusion body myositis: implications for clinical trials. *Neurology* 2001;57:548–50.
8 Benveniste O, Guiguet M, Freebody J, *et al*. Long-term observational study of sporadic inclusion body myositis. *Brain* 2011;134(Pt 11):3176–84.
9 Griggs RC, Askanas V, DiMauro S, *et al*. Inclusion body myositis and myopathies. *Ann Neurol* 1995;38:705–13.
10 Hilton-Jones D, Miller A, Parton M, *et al*. Inclusion body myositis: MRC Centre for Neuromuscular Diseases, IBM workshop, London, 13 June 2008. *Neuromuscul Disord* 2010;20:142–7.
11 Verschuuren JJ, Badrising UA, Van Engelen BG, *et al*. Inclusion body myositis. In: Emery AEH, ed. *Diagnostic criteria for neuromuscular disorders*. London: Royal Society of Medicine Press, 1997:81–4.
12 Brady S, Squier W, Hilton-Jones D. Clinical assessment determines the diagnosis of inclusion body myositis independently of pathological features. *J Neurol Neurosurg Psychiatry* 2013;84:1240–6.
13 Felice KJ, North WA. Inclusion body myositis in Connecticut: observations in 35 patients during an 8-year period. *Medicine (Baltimore)* 2001;80:320–7.
14 Brady S, Squier W, Sewry C, *et al*. A retrospective cohort study identifying the principal pathological features useful in the diagnosis of inclusion body myositis. *BMJ Open* 2014;4:e004552.
15 Cortese A, Machado P, Morrow J, *et al*. Longitudinal observational study of sporadic inclusion body myositis: implications for clinical trials. *Neuromuscul Disord* 2013;23:404–12.
16 Pluk H, van Hoeve BJ, van Dooren SH, *et al*. Autoantibodies to cytosolic 5′-nucleotidase 1A in inclusion body myositis. *Ann Neurol* 2013;73:397–407.
17 Salajegheh M, Lam T, Greenberg SA. Autoantibodies against a 43 KDa muscle protein in inclusion body myositis. *PLoS ONE* 2011;6:e20266.
18 Goyal NA, Cash TM, Alam U, *et al*. Seropositivity for NT5c1A antibody in sporadic inclusion body myositis predicts more severe motor, bulbar and respiratory involvement. *J Neurol Neurosurg Psychiatry* 2016;87:373–8.
19 Larman HB, Salajegheh M, Nazareno R, *et al*. Cytosolic 5′-nucleotidase 1A autoimmunity in sporadic inclusion body myositis. *Ann Neurol* 2013;73:408–18.
20 Rothwell S, Cooper RG, Lundberg IE, *et al*. Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups. *Ann Rheum Dis* 2016;75:1558–66.
21 Lundberg IE, Svensson J. Registries in idiopathic inflammatory myopathies. *Curr Opin Rheumatol* 2013;25:729–34.
22 Herbert MK, Stammen-Vogelzangs J, Verbeek MM, *et al*. Disease specificity of autoantibodies to cytosolic 5′-nucleotidase 1A in sporadic inclusion body myositis versus known autoimmune diseases. *Ann Rheum Dis* 2016;75:696–701.
23 Tansley SL, Betteridge ZE, Gunawardena H, *et al*. Anti-MDA5 autoantibodies in juvenile dermatomyositis identify a distinct clinical phenotype: a prospective cohort study. *Arthritis Res Ther* 2014;16:R138.
24 Lloyd TE, Christopher-Stine L, Pinal-Fernandez I, *et al*. Cytosolic 5′-nucleotidase 1A as a target of circulating autoantibodies in autoimmune diseases. *Arthritis Care Res* 2016;68:66–71.
25 Limaye VS, Lester S, Blumbergs P, *et al*. Anti-C N1A antibodies in South Australian patients with inclusion body myositis. *Muscle Nerve* 2016;53:654–5.
26 Herbert MK, Pruijn GJ. Novel serology testing for sporadic inclusion body myositis: disease-specificity and diagnostic utility. *Curr Opin Rheumatol* 2015;27:595–600.

OPEN ACCESS

EXTENDED REPORT

GWAS of clinically defined gout and subtypes identifies multiple susceptibility loci that include urate transporter genes

Akiyoshi Nakayama,[1] Hirofumi Nakaoka,[2] Ken Yamamoto,[3] Masayuki Sakiyama,[1,4] Amara Shaukat,[5] Yu Toyoda,[6] Yukinori Okada,[7,8,9] Yoichiro Kamatani,[8] Takahiro Nakamura,[10] Tappei Takada,[6] Katsuhisa Inoue,[11] Tomoya Yasujima,[12] Hiroaki Yuasa,[12] Yuko Shirahama,[3] Hiroshi Nakashima,[13] Seiko Shimizu,[1] Toshihide Higashino,[1] Yusuke Kawamura,[1] Hiraku Ogata,[1] Makoto Kawaguchi,[1] Yasuyuki Ohkawa,[14] Inaho Danjoh,[15] Atsumi Tokumasu,[16] Keiko Ooyama,[16] Toshimitsu Ito,[17] Takaaki Kondo,[18] Kenji Wakai,[19] Blanka Stiburkova,[20,21] Karel Pavelka,[21] Lisa K Stamp,[22] Nicola Dalbeth,[23] Eurogout Consortium Yutaka Sakurai,[13] Hiroshi Suzuki,[6] Makoto Hosoyamada,[24] Shin Fujimori,[25] Takashi Yokoo,[26] Tatsuo Hosoya,[26,27] Ituro Inoue,[2] Atsushi Takahashi,[8,28] Michiaki Kubo,[29] Hiroshi Ooyama,[16] Toru Shimizu,[30,31] Kimiyoshi Ichida,[27,32] Nariyoshi Shinomiya,[1] Tony R Merriman,[5] Hirotaka Matsuo[1]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-209632).

For numbered affiliations see end of article.

Correspondence to
Dr Hirotaka Matsuo, Department of Integrative Physiology and Bio-Nano Medicine, National Defense Medical College, 3-2 Namiki, Tokorozawa, Saitama 359-8513, Japan; hmatsuo@ndmc.ac.jp

AN, HNakaoka, KY, MS, AS and YT contributed equally.

Received 29 March 2016
Revised 4 November 2016
Accepted 5 November 2016
Published Online First 29 November 2016

To cite: Nakayama A, Nakaoka H, Yamamoto K, *et al*. *Ann Rheum Dis* 2017;**76**:869–877.

ABSTRACT

Objective A genome-wide association study (GWAS) of gout and its subtypes was performed to identify novel gout loci, including those that are subtype-specific.

Methods Putative causal association signals from a GWAS of 945 clinically defined gout cases and 1213 controls from Japanese males were replicated with 1396 cases and 1268 controls using a custom chip of 1961 single nucleotide polymorphisms (SNPs). We also first conducted GWASs of gout subtypes. Replication with Caucasian and New Zealand Polynesian samples was done to further validate the loci identified in this study.

Results In addition to the five loci we reported previously, further susceptibility loci were identified at a genome-wide significance level ($p<5.0\times10^{-8}$): urate transporter genes (*SLC22A12* and *SLC17A1*) and *HIST1H2BF-HIST1H4E* for all gout cases, and *NIPAL1* and *FAM35A* for the renal underexcretion gout subtype. While *NIPAL1* encodes a magnesium transporter, functional analysis did not detect urate transport via NIPAL1, suggesting an indirect association with urate handling. Localisation analysis in the human kidney revealed expression of NIPAL1 and FAM35A mainly in the distal tubules, which suggests the involvement of the distal nephron in urate handling in humans. Clinically ascertained male patients with gout and controls of Caucasian and Polynesian ancestries were also genotyped, and *FAM35A* was associated with gout in all cases. A meta-analysis of the three populations revealed *FAM35A* to be associated with gout at a genome-wide level of significance ($p_{meta}=3.58\times10^{-8}$).

Conclusions Our findings including novel gout risk loci provide further understanding of the molecular pathogenesis of gout and lead to a novel concept for the therapeutic target of gout/hyperuricaemia.

INTRODUCTION

Gout is a common disease characterised by acute painful arthritis, and its global burden continues to rise with the increasingly ageing population.[1] Gout is caused by hyperuricaemia, and can be classified according to patients' clinical parameters reflecting its causes[2 3] as renal overload (ROL) gout and renal underexcretion (RUE) gout. As shown in online supplementary figure S1, patients with gout with increased urinary excretion of urate due to overproduction and/or decreased extra-renal underexcretion of urate are classified as having ROL gout, whereas those with decreased renal excretion of urate are defined as having RUE gout.[2] Reflecting their causes, almost all patients with gout are divided into those two subtypes. Although these subtypes are important from both genetic and pathophysiological points of view,[2 4] genome-wide association studies (GWASs) of gout subtypes have never been performed, partly due to the difficulty in assembling sufficient gout cases with requisite clinical data, including data from a time-consuming urinary collection examination.

We and other groups[5–9] recently reported gout/hyperuricaemia to have relatively strong genetic risk factors. More recently, and for the first time, we performed a GWAS with only clinically defined Japanese male gout cases in which 16 single nucleotide polymorphisms (SNPs) were replicated, and five gout-risk loci were identified including two novel loci (*MYL2-CUX2* and *CNIH-2*).[10] In the present study (see online supplementary figure S2), we extended our analysis to identify novel susceptibility loci for gout by replicating approximately 2000 SNPs top-ranked in the GWASs of all gout and/or its subtypes. In addition, for the first time, we performed GWASs of gout subtypes to identify

subtype-specific (cause-specific) risk loci. Furthermore, we conducted a replication study with independent Caucasian and Polynesian populations to validate loci.

METHODS

Subjects and genotyping

Genome-wide genotyping was performed with the Illumina HumanOmniExpress-12 v1.0 (Illumina) platform using 946 clinically defined gout cases and 1213 controls, all Japanese males. Detailed methods of genotyping and quality control are previously described.[10] Ultimately, 570 442 SNPs passed filters for 945 cases and 1213 controls. At the replication stage, 1246 cases were genotyped with a custom genotype platform using iSelect HD Custom Genotyping BeadChips (Illumina) on 1961 SNPs, as described in online supplementary methods and supplementary figure S3, and 150 gout cases were genotyped with the Illumina HumanOmniExpress-24 v1.0 (Illumina) platform. For controls, 1268 Japanese males with a serum uric acid (SUA) level ≤ 7.0 mg/dL and without gout history were recruited from BioBank Japan[11 12] and genotyped with the Illumina HumanOmniExpress-12 v1.0 (Illumina) platform. Finally, 1961 SNPs with 1396 gout cases and 1268 controls were successfully genotyped (see online supplementary table S1). A genome-wide significance threshold was set to be $\alpha=5.0\times10^{-8}$ to claim evidence of a significant association.

GWASs of the two subtypes of gout, ROL gout and RUE gout (see online supplementary figure S1), were also performed, followed by replication studies with a custom SNP chip (see online supplementary figure S3) and a subsequent meta-analysis. As described previously,[2 10] and shown in online supplementary figure S1 and supplementary methods, ROL gout and RUE gout are defined when patients' urinary urate excretion is over 25.0 mg/hour/1.73 m^2 (600 mg/day/1.73 m^2) and patients' urate clearance (urate clearance/creatinine clearance ratio, FE_{UA}) is under 5.5%, respectively. For GWASs of gout subtypes, 1178 cases were classified as ROL gout (560 cases at GWAS stage and 618 cases at replication stage) and 1315 cases as RUE gout (619 cases at GWAS stage and 696 cases at replication stage), respectively (see online supplementary table S2).

A replication study with independent Caucasian and New Zealand (NZ) Polynesian sample sets was also performed to validate the genetic risk loci identified in the present study. This replication was done in a data set recruited from New Zealand[13] and from Europe by the Eurogout Consortium[14] comprising 1319 male cases and 514 male controls of European ancestry and 971 male cases and 565 male controls of NZ Polynesian ancestry. SNPs were genotyped by an allelic discrimination assay (TaqMan) with a LightCycler 480 Real-Time PCR (RT-PCR) System (Roche Applied Science, Indianapolis, Indiana, USA). Detailed information of clinical characteristics and genetic analysis is shown in online supplementary methods and tables S1–S3.

Statistical analyses

The inverse-variance fixed-effects model was used for meta-analysis. In the meta-analysis with Japanese, Caucasian and NZ Polynesian populations or in the presence of heterogeneity ($p_{het} < 0.05$ or $I^2 > 50\%$), we implemented the DerSimonian and Laird random-effects model for meta-analysis.[15] For the replication analysis with Caucasian and NZ Polynesian sample sets, ORs were adjusted by age and ancestral group. All the meta-analyses were performed using the R V.3.1.1 and 3.2.2 (R Development Core Team. R: a language and environment for statistical computing. Vienna: R. Foundation for Statistical Computing, 2006) with meta package. All calculations of linkage disequilibrium (LD, measured in r^2) were conducted using the Japanese population. The detailed information for statistical analyses is described in online supplementary methods.

Functional and localisation analyses

Urate transport analysis of NIPAL1 was performed with an oocyte expression system[16 17] with high potassium (HK) buffer or HK buffer without magnesium. For immunohistochemical analysis, the human kidney sections (3 μm) incubated with anti-human NIPAL1 antibody (1:50) (LS-C164878; LifeSpan BioSciences, Washington, USA) or with anti-human FAM35A antibody (1:75) (HPA036582; Sigma-Aldrich, Missouri, USA) were used, and then visualised with diaminobenzidine (0.8 mM).[18 19] Intracellular localisation of NIPAL1 was also studied in *Xenopus* oocytes and Madin-Darby canine kidney II (MDCKII) cells. Detailed information for the functional and localisation analyses is described in online supplementary methods.

RESULTS

GWAS of all gout and its subtypes

In addition to the GWAS stage previously performed with 945 patients with clinically defined gout and 1213 controls, all Japanese males[10] (see online supplementary figure S4), the replication stage for all cases of gout was carried out by genotyping 1961 SNPs (see online supplementary figure S3 and supplementary note) in a further 1396 male patients and 1268 male controls, and a meta-analysis then conducted (see online supplementary figure S2). Furthermore, GWASs of two subtypes of gout, ROL gout (figure 1A) and RUE gout (figure 1B), were also performed in the present study, followed by replication studies with a custom SNP chip and a subsequent meta-analysis.

Meta-analysis of both the GWAS and the replication study for all gout cases (table 1) identified eight loci which showed evidence for associations at the genome-wide significance level: rs3114020 of *ABCG2* ($p_{meta}=8.66\times10^{-35}$; OR=1.89), rs1014290 of *SLC2A9* ($p_{meta}=6.50\times10^{-26}$; OR=1.57), rs4766566 of *CUX2* ($p_{meta}=4.03\times10^{-20}$; OR=1.51), rs2285340 of *SLC22A12* ($p_{meta}=4.61\times10^{-11}$; OR=1.40), rs1260326 of *GCKR* ($p_{meta}=7.19\times10^{-11}$; OR=1.31), rs1165176 of *SLC17A1* ($p_{meta}=1.47\times10^{-9}$; OR=1.42), rs11758351 of *HIST1H2BF-HIST1H4E* ($p_{meta}=1.63\times10^{-8}$; OR=1.40) and rs4073582 of *CNIH-2* ($p_{meta}=3.56\times10^{-8}$; OR=1.58). Among these eight loci, *SLC22A12*, *SLC17A1* and *HIST1H2BF-HIST1H4E* (figure 2A–C) were first identified as gout-risk loci by the GWAS approach at the genome-wide significance level. *SLC17A1* was identified here by the GWAS approach for the first time, while Hollis-Moffatt *et al*[20] reported that rs1183201, another SNP of *SLC17A1*, is strongly associated with gout in Caucasians and NZ Polynesian sample sets by the candidate gene approach. While rs11758351 of *HIST1H2BF-HIST1H4E* is located 374 kb downstream from rs1165176 of *SLC17A1*, they are not in LD with each other ($r^2=0.03$), demonstrating them to be independent susceptibility loci for gout. There was also a significant signal from rs2532941 of *VARS2* ($p_{meta}=2.74\times10^{-8}$; OR=1.32), which is located downstream of *HIST1H2BF-HIST1H4E* by 4.7 Mb, and is reported to be associated with mitochondrial respiration.[21] Since rs2532941 of *VARS2* showed mild LD with rs11758351 of *HIST1H2BF-HIST1H4E* ($r^2=0.37$), its significance did not remain for the GWAS stage samples after adjustment with rs11758351 of *HIST1H2BF-HIST1H4E* (p=0.08), or with both rs1165176 of *SLC17A1* and rs11758351 (p=0.11).

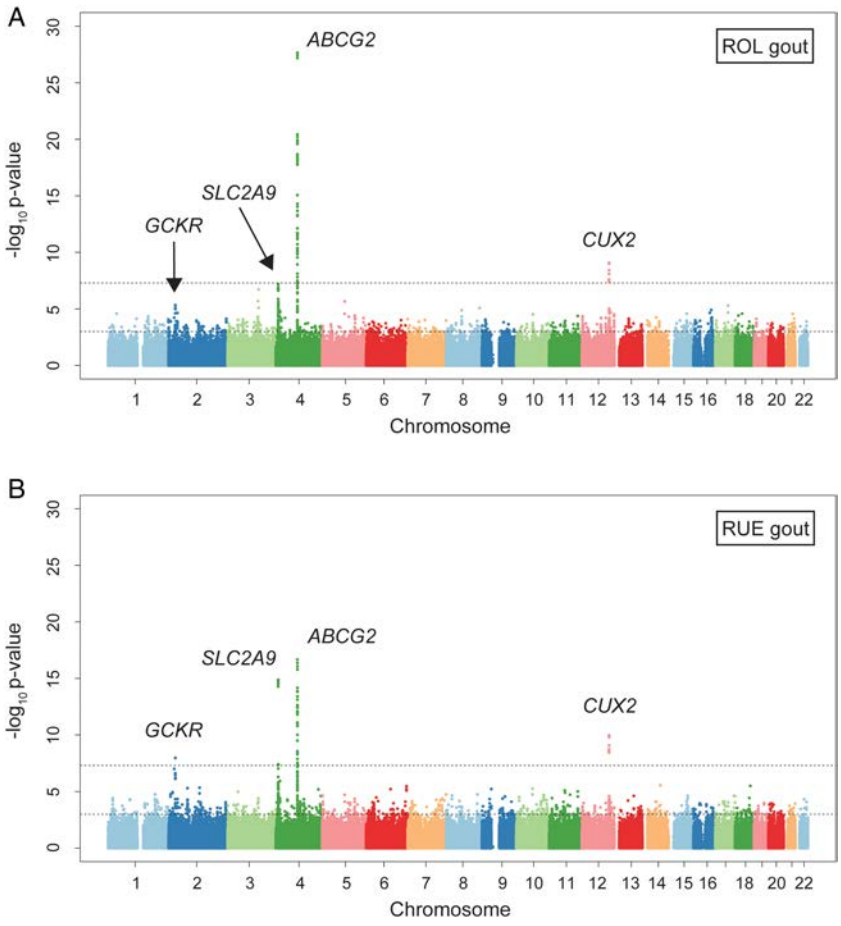


Figure 1 Manhattan plots of GWASs of subtypes of gout. Manhattan plots of GWASs of (A) ROL gout subtype and (B) RUE gout subtype. X-axis shows chromosomal positions. Y-axis shows $-\log_{10}$ p values. The upper and lower dotted lines indicate the genome-wide significance threshold ($p=5.0\times10^{-8}$) and the cut-off level for selecting single nucleotide polymorphisms for replication study ($p=0.001$), respectively. GWAS, genome-wide association study; ROL, renal overload; RUE, renal underexcretion.

For GWASs of gout subtypes, 1178 cases were classified as ROL gout (560 cases at GWAS stage and 618 cases at replication stage) and 1315 cases as RUE gout (619 cases at GWAS stage and 696 cases at replication stage), respectively (see online supplementary table S2). The meta-analysis of a GWAS of the ROL gout subtype and a replication study revealed significant SNPs in the following four loci: rs2728104 of *ABCG2* ($p_{meta}=5.08\times10^{-33}$; OR=1.84), rs4766566 of *CUX2* ($p_{meta}=8.14\times10^{-17}$; OR=1.59), rs3733589 of *SLC2A9* ($p_{meta}=2.25\times10^{-13}$; OR=1.47) and rs1260326 of *GCKR* ($p_{meta}=5.39\times10^{-9}$; OR=1.35).

Another subtype analysis, that is, the meta-analysis of a GWAS of RUE gout and a replication study (table 1) demonstrated significant SNPs in the following seven loci: rs1014290 of *SLC2A9* ($p_{meta}=8.71\times10^{-25}$; OR=1.69), rs1871744 of *ABCG2* ($p_{meta}=2.49\times10^{-22}$; OR=1.81), rs4766566 of *CUX2* ($p_{meta}=2.17\times10^{-18}$; OR=1.60), rs2285340 of *SLC22A12* ($p_{meta}=8.79\times10^{-10}$; OR=1.44), rs780094 of *GCKR* ($p_{meta}=1.62\times10^{-9}$; OR=1.35), rs11733284 of *NIPAL1* ($p_{meta}=1.13\times10^{-8}$; OR=1.34) and rs7903456 of *FAM35A* ($p_{meta}=4.29\times10^{-8}$; OR=1.34). The latter two loci, *NIPAL1* and *FAM35A*, were novel risk loci by the GWAS of the RUE gout subtype (figure 2D, E). In total, 10 loci were identified from the present GWAS of gout and its subtypes (table 1 and see online supplementary table S4).

Of the seven loci newly identified by GWAS of the RUE gout subtype, only *NIPAL1* and *FAM35A* had not been implicated previously in the GWASs of SUA levels or gout. Analysis with data from previously reported GWAS[22] of SUA in Caucasians revealed the association with *NIPAL1* and *FAM35A* loci (see online supplementary figure S5).

Urate transport analysis of NIPAL1 transporter

NIPAL1 and *FAM35A* were revealed to be associated with RUE gout in the present study. NIPAL1 has been reported to be a magnesium transporter,[23] which has nine transmembrane domains (figure 3A), whereas FAM35A is predicted to be a soluble protein. In this context, we hypothesised that NIPAL1 could be involved in the regulation of urate handling as a renal urate efflux transporter. However, our functional analysis using *Xenopus* oocytes did not show urate transport via NIPAL1, regardless of the presence of magnesium (figure 3B).

Localisation analysis of NIPAL1 and FAM35A

By immunohistochemical analysis, NIPAL1 and FAM35A showed cytosolic expression in the renal distal tubules of human kidney (figure 4A, B). Both proteins were also weakly detected in the cytoplasm of collecting ducts. NIPAL1-expressing *Xenopus* oocytes and MDCKII cells also showed intracellular localisation of NIPAL1 (see online supplementary figure S6).

Replication study of all gout cases with Caucasian and Polynesian populations

A replication study for the discovered loci (*SLC22A12*, *SLC17A1*, *HIST1H2BF-HIST1H4E*, *NIPAL1* and *FAM35A*) was performed for all gout cases with males drawn from Caucasian (1319 cases and 514 controls) and NZ Polynesian populations (971 cases and 565 controls). Because a gain-of-function SNP of *SLC17A1*, rs1165196 (Ile269Thr),[16] was in strong LD with rs1165176 ($r^2=0.99$), we performed the following analyses using rs1165196, assuming that the causal SNP in this locus was rs1165196 of *SLC17A1*. Among these five loci, the

Table 1 Single nucleotide polymorphisms (SNPs) associated with gout and its subtypes at a genome-wide level of significance in the Japanese population

						GWAS¶				Replication study**				Meta-analysis††			
			Position			Frequency of A1				Frequency of A1						Heterogeneity	
Gout types	SNP*	Chr.	(bp)†	Gene‡	A1/A2§	Cases	Controls	OR (95% CI)	p Value	Cases	Controls	OR (95% CI)	p Value	OR (95% CI)	p Value	Cochran's Q	I^2 (%)
All gout	rs1260326	2	27730940	*GCKR*	T/C	0.616	0.535	1.39 (1.23 to 1.57)	1.34×10^{-7}	0.611	0.557	1.25 (1.12 to 1.39)	6.10×10^{-5}	1.31 (1.21 to 1.42)	7.19×10^{-11}	0.20	38.2
	rs1014290	4	10001861	*SLC2A9*	T/C	0.678	0.564	1.63 (1.44 to 1.85)	1.75×10^{-14}	0.673	0.576	1.51 (1.35 to 1.69)	2.97×10^{-13}	1.57 (1.44 to 1.70)	6.50×10^{-26}	0.39	0.0
	rs3114020	4	89083666	*ABCG2*	C/T	0.842	0.724	2.03 (1.75 to 2.37)	1.17×10^{-20}	0.844	0.752	1.78 (1.55 to 2.04)	7.74×10^{-17}	1.89 (1.71 to 2.09)	8.66×10^{-35}	0.20	38.9
	rs1165176	6	25830298	***SLC17A1***	G/A	0.874	0.834	1.38 (1.16 to 1.64)	2.89×10^{-4}	0.872	0.824	1.46 (1.25 to 1.69)	1.08×10^{-6}	1.42 (1.27 to 1.59)	1.47×10^{-9}	0.63	0.0
	rs11758351	6	26203910	***HIST1H2BF-HIST1H4E***	G/T	0.158	0.121	1.37 (1.15 to 1.63)	4.22×10^{-4}	0.158	0.116	1.43 (1.22 to 1.67)	1.01×10^{-5}	1.40 (1.25 to 1.57)	1.63×10^{-8}	0.72	0.0
	rs2285340	11	64435906	***SLC22A12***	A/G	0.228	0.174	1.40 (1.21 to 1.63)	1.09×10^{-5}	0.227	0.174	1.40 (1.22 to 1.61)	9.96×10^{-7}	1.40 (1.27 to 1.55)	4.61×10^{-11}	1.00	0.0
	rs4073582	11	66050712	*CNIH-2*	C/T	0.950	0.915	1.78 (1.39 to 2.29)	4.32×10^{-6}	0.943	0.920	1.44 (1.16 to 1.79)	8.47×10^{-4}	1.58 (1.34 to 1.86)	3.56×10^{-8}	0.21	36.1
	rs4766566	12	111706877	*CUX2*	T/C	0.735	0.633	1.60 (1.41 to 1.83)	1.22×10^{-12}	0.741	0.665	1.44 (1.28 to 1.62)	2.07×10^{-9}	1.51 (1.38 to 1.65)	4.03×10^{-20}	0.22	33.8
ROL gout	rs1260326	2	27730940	*GCKR*	T/C	0.611	0.535	1.36 (1.18 to 1.58)	2.43×10^{-5}	0.626	0.557	1.33 (1.16 to 1.53)	6.12×10^{-5}	1.35 (1.22 to 1.49)	5.39×10^{-9}	0.81	0.0
	rs3733589	4	9987324	*SLC2A9*	G/A	0.662	0.570	1.48 (1.28 to 1.71)	2.00×10^{-7}	0.668	0.580	1.46 (1.26 to 1.68)	2.05×10^{-7}	1.47 (1.32 to 1.63)	2.25×10^{-13}	0.88	0.0
	rs2728104	4	88973006	*ABCG2*	C/T	0.505	0.346	1.93 (1.67 to 2.23)	3.28×10^{-19}	0.496	0.359	1.75 (1.53 to 2.01)	1.56×10^{-15}	1.84 (1.66 to 2.03)	5.08×10^{-33}	0.35	0.0
	rs4766566	12	111706877	*CUX2*	T/C	0.737	0.633	1.62 (1.39 to 1.90)	8.42×10^{-10}	0.757	0.665	1.57 (1.34 to 1.83)	7.55×10^{-9}	1.59 (1.43 to 1.78)	8.14×10^{-17}	0.76	0.0
RUE gout	rs780094	2	27741237	*GCKR*	T/C	0.633	0.543	1.45 (1.26 to 1.67)	2.43×10^{-7}	0.615	0.559	1.26 (1.10 to 1.44)	6.47×10^{-4}	1.35 (1.22 to 1.48)	1.62×10^{-9}	0.16	48.8
	rs1014290	4	10001861	*SLC2A9*	T/C	0.699	0.564	1.80 (1.55 to 2.08)	1.58×10^{-15}	0.685	0.576	1.60 (1.39 to 1.84)	1.72×10^{-11}	1.69 (1.53 to 1.87)	8.71×10^{-25}	0.26	21.8
	rs11733284	4	48028097	***NIPAL1***	A/G	0.346	0.281	1.35 (1.17 to 1.57)	6.48×10^{-5}	0.342	0.280	1.34 (1.16 to 1.54)	6.36×10^{-5}	1.34 (1.21 to 1.49)	1.13×10^{-8}	0.91	0.0
	rs1871744	4	89039629	*ABCG2*	T/C	0.834	0.723	1.93 (1.62 to 2.29)	3.85×10^{-14}	0.824	0.733	1.71 (1.45 to 2.01)	7.04×10^{-11}	1.81 (1.60 to 2.04)	2.49×10^{-22}	0.33	0.0
	rs7903456	10	88919319	***FAM35A***	A/G	0.303	0.248	1.32 (1.13 to 1.53)	4.32×10^{-4}	0.296	0.235	1.37 (1.18 to 1.59)	3.09×10^{-5}	1.34 (1.21 to 1.49)	4.29×10^{-8}	0.72	0.0
	rs2285340	11	64435906	***SLC22A12***	A/G	0.236	0.174	1.47 (1.25 to 1.74)	8.04×10^{-6}	0.228	0.174	1.41 (1.20 to 1.66)	4.04×10^{-5}	1.44 (1.28 to 1.62)	8.79×10^{-10}	0.72	0.0
	rs4766566	12	111706877	*CUX2*	T/C	0.738	0.633	1.63 (1.40 to 1.89)	1.58×10^{-10}	0.759	0.665	1.58 (1.36 to 1.83)	9.51×10^{-10}	1.60 (1.44 to 1.78)	2.17×10^{-18}	0.78	0.0

*dbSNP rs number. SNPs having associations for all gout, ROL gout and RUE gout at the lowest p value in each locus by meta-analysis are shown in this table.
†SNP positions are based on NCBI human genome reference sequence build 37.4.
‡Five discovered loci are shown in bold.
§A1 is risk-associated allele and A2 is non-risk-associated allele.
¶945 cases for all gout, 560 cases for ROL gout, 619 cases for RUE gout with 1213 controls from Japanese male population.
**1396 cases for all gout, 618 cases for ROL gout, 696 cases for RUE gout with 1268 controls from Japanese male population.
††Meta-analysis of GWAS and replication samples.
Chr., chromosome; GWAS, genome-wide association study; ROL, renal overload; RUE, renal underexcretion; SNP, single nucleotide polymorphism.

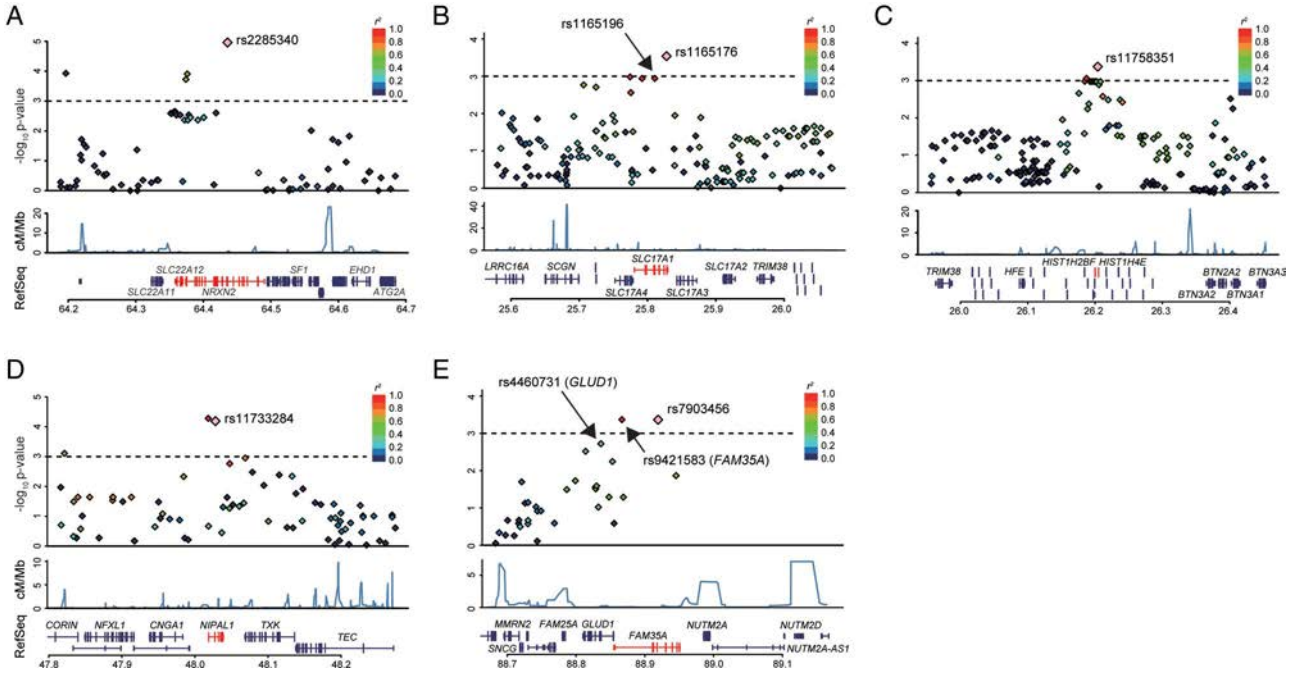


Figure 2 Regional association plots of five discovered loci. Three loci were revealed to exceed the genome-wide significance level from the meta-analysis with all gout cases, and two loci with renal underexcretion (RUE) gout cases. The highest association signal in each panel is located on (A) *SLC22A12*, (B) *SLC17A1* and (C) *HIST1H2BF-HIST1H4E* for all gout cases, and (D) *NIPAL1* and (E) *FAM35A* for RUE gout cases. The region within 250 kb from the single nucleotide polymorphism (SNP) indicating the lowest p value is shown. (Top panel) Plots of $-\log_{10}$ p values for the test of SNP association with gout in the genome-wide association study stage. The SNP showing the lowest p value in the meta-analysis is depicted as a pink diamond. Other SNPs are colour-coded according to the extent of linkage disequilibrium (measured in r^2) with the SNP showing the lowest p value. (Middle panel) Recombination rates (centimorgans per Mb) estimated from HapMap Phase II data are plotted. (Bottom panel) RefSeq genes. Genomic coordinates are based on NCBI human genome reference sequence build 37.

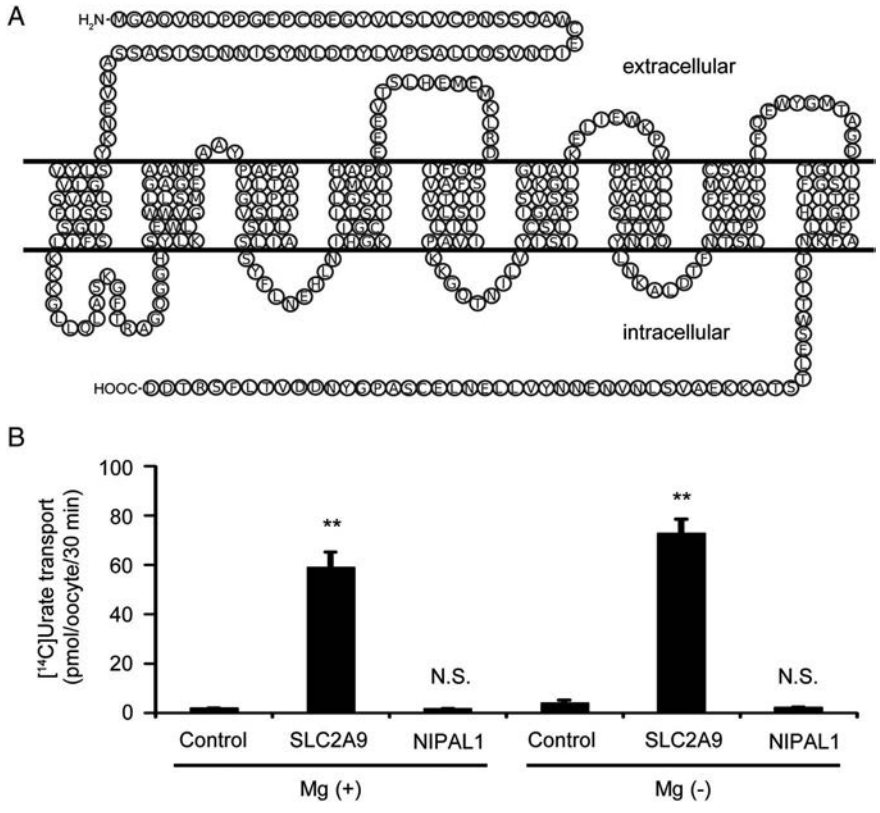


Figure 3 Functional analysis of NIPAL1 transporter. (A) The topological model of the NIPAL1 transporter. NIPAL1 is predicted to have nine transmembrane regions. The amino acid sequences of NIPAL1 were obtained from GenBank (accession code NM_207330). (B) Urate transport analysis of NIPAL1. SLC2A9 (also known as GLUT9) is a renal urate transporter and is used for a positive control for the urate transport analysis. In contrast to SLC2A9, urate transport via NIPAL1 was not detected, regardless of the presence of magnesium. Data are expressed as mean±SEM (n=8). Statistical analyses for significant differences were performed according to Student's t-test. (**p<0.01; N.S., not significantly different as compared with control.).

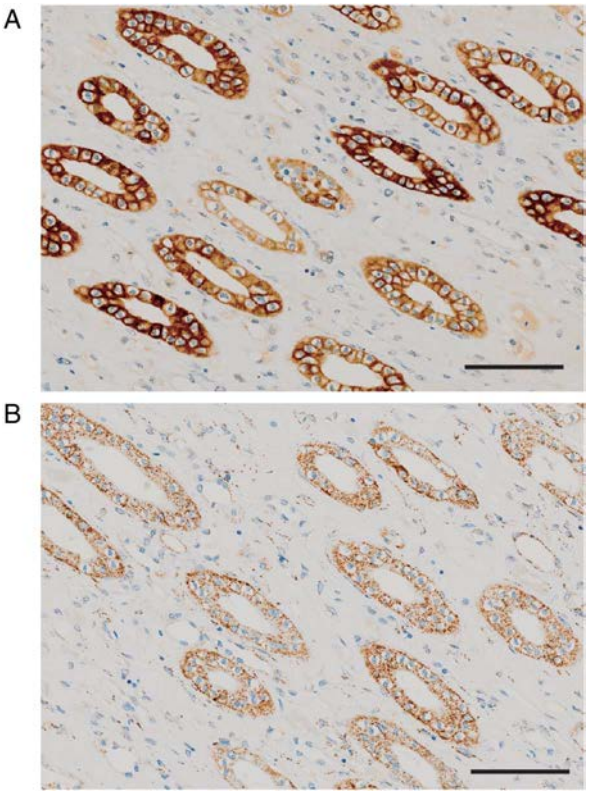


Figure 4 Localisation analysis of NIPAL1 and FAM35A in the human kidney. Cytosolic expression was detected strongly in distal tubules and weakly in collecting ducts in human kidney for (A) NIPAL1 protein and (B) FAM35A protein. Bar=100 μm.

meta-analysis of those populations for all gout revealed a significant association with rs7903456 of *FAM35A* ($p_{meta}=9.72\times10^{-3}$; OR=1.17) (table 2). Although *SLC17A1* did not show significance ($p_{meta}=0.119$) in the present study of those populations (table 2), a previous paper[20] revealed a significant association of *SLC17A1* with gout in Caucasian and NZ Polynesian sample sets, indicating the necessity of further replication studies to investigate the ancestral differences in the significance of other genetic loci including *SLC17A1*. Genotyping the *CUX2* and *CNIH-2* loci, which were identified in both our present and previous GWASs of Japanese,[10] was also performed, and the *CUX2* locus was replicated successfully for the first time in other populations (see online supplementary table S5). The results of further association analyses and expression quantitative trait locus (eQTL) analysis are shown in online supplementary note and tables S6 and S7. Significant effects on FE_{UA} were detected in *NIPAL1*, *FAM35A* and *SLC22A12* loci in the Japanese population, and were also observed at *SLC17A1* in NZ Polynesian population.

A further meta-analysis of all gout cases with Japanese, Caucasian and NZ Polynesian populations was performed for *NIPAL1* and *FAM35A*, which were at a genome-wide significance level in the Japanese population only for the RUE gout subtype, and not for all gout cases. rs11733284 of *NIPAL1* was not associated with all gout ($p_{meta}=0.16$; OR=1.11), suggesting the presence of ancestral differences in genetic effects of this locus, or a subtype-specific effect. On the other hand, rs7903456 of *FAM35A* showed an association with all gout at a genome-wide level of significance ($p_{meta}=3.58\times10^{-8}$; OR=1.23) (figure 5), indicating that rs7903456 is a susceptibility locus for all gout as well as the RUE gout subtype.

Table 2 Replication study of all gout for five discovered loci in Caucasian and NZ Polynesian sample sets

SNP*	Chr..	position (bp)†	Gene	A1/A2‡	Caucasian§				NZ Polynesian¶				Meta-analysis**		Heterogeneity	
					Frequency of A1		OR (95% CI)	p Value	Frequency of A1		OR (95% CI)	p Value	OR (95% CI)	p Value	Cochran's Q	I^2 (%)
					Cases	Controls			Cases	Controls						
rs11733284	4	48028097	*NIPAL1*	A/G	0.362	0.356	1.01 (0.86 to 1.18)	0.896	0.251	0.270	0.92 (0.77 to 1.10)	0.355	0.97 (0.86 to 1.09)	0.603	0.43	0.0
rs1165196	6	25813150	*SLC17A1*	T/C	0.614	0.583	1.11 (0.95 to 1.30)	0.271	0.731	0.711	1.12 (0.93 to 1.35)	0.266	1.11 (0.98 to 1.25)	0.119	0.88	0.0
rs11758351	6	26203910	*HIST1H2BF-HIST1H4E*	G/T	0.141	0.158	0.86 (0.70 to 1.07)	0.173	0.192	0.199	0.90 (0.74 to 1.10)	0.334	0.88 (0.77 to 1.02)	0.0941	0.77	0.0
rs7903456	10	88919319	*FAM35A*	A/G	0.737	0.699	1.18 (1.00 to 1.40)	0.0462	0.351	0.333	1.16 (0.98 to 1.38)	0.0997	1.17 (1.04 to 1.32)	9.72×10^{-3}	0.85	0.0
rs2285340††	11	64435906	*SLC22A12*	A/G	–	–	–	–	0.158	0.143	1.06 (0.84 to 1.35)	0.634	–	–		

*dbSNP rs number.
†SNP positions are based on NCBI human genome reference sequence build 37.4.
‡A1 is risk-associated allele, and A2 is non-risk-associated allele.
§1319 cases for all gout and 514 controls from Caucasian male population.
¶971 cases for all gout and 565 controls from NZ Polynesian male population.
**Meta-analysis of Caucasian and NZ Polynesian samples.
††rs2285340 is monomorphic in Caucasians.
Chr., chromosome; NZ, New Zealand; SNP, single nucleotide polymorphism.

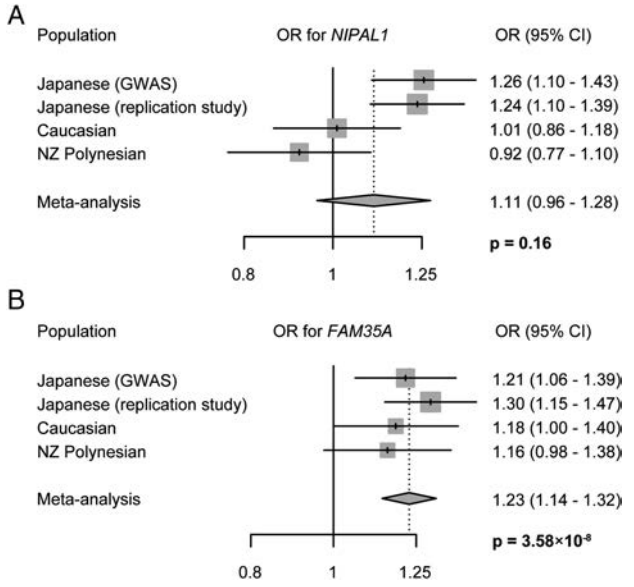


Figure 5 Forest plots for all gout among Japanese, Caucasian and New Zealand (NZ) Polynesian populations. Although rs11733284 of *NIPAL1* (A) did not show significant association with all gout, rs7903456 of *FAM35A* (B) revealed an association with all gout at a genome-wide significance level (p_{meta}=3.58×10^{-8}; OR=1.23). GWAS, genome-wide association study.

Meta-analysis of all gout for the other three loci (*SLC22A12*, *SLC17A1* and *HIST1H2BF-HIST1H4E*) was also performed with Japanese, Caucasian and NZ Polynesian populations as shown in online supplementary figure S7. rs11758351 of *HIST1H2BF-HIST1H4E* did not show a significant association with gout (p_{meta}=0.40; OR=1.12). rs2285340 of *SLC22A12* and rs1165196 of *SLC17A1* did not reach a genome-wide level of significance (p_{meta}=2.47×10^{-4}; OR=1.31; and p_{meta}=1.28×10^{-3}; OR=1.25, respectively) partly due to statistical fluctuation in relatively small sample sets, although the effects were consistently in the same direction.

DISCUSSION

With clinically defined gout cases, we previously performed a GWAS[10] and revealed that *ABCG2*, *SLC2A9*, *MYL2-CUX2*, *GCKR* and *CNIH-2* were associated with gout at a genome-wide significance level (see online supplementary figure S4). A more recent GWAS by Li *et al*[24] with clinically ascertained gout cases revealed three novel loci (*BCAS3*, *RFX3* and *KCNQ1*) in Han Chinese. In the present study, we performed a gout follow-up study focused on loci not reaching the genome-wide level of significance in the previous GWAS,[10] genotyping 1961 SNPs in an additional 1396 cases and 1268 controls. We revealed a total of eight loci to be associated with all gout cases in Japanese males (table 1). Among them, three loci (*SLC22A12*, *SLC17A1* and *HIST1H2BF-HIST1H4E*) were first identified as gout risk loci at a genome-wide significance level by the present GWAS approach.

Both *SLC22A12* and *SLC17A1* encode urate transporters at the apical side of the renal proximal tubule[16 25] (see online supplementary figure S8) and are reportedly associated with SUA level in humans by previous GWASs of SUA.[12 22 26 27] Therefore, it is reasonable that SNPs around these loci would display significant associations with gout or sequelae of hyperuricaemia (see also online supplementary note for detail).

The *HIST1H2BF* and *HIST1H4E* genes encode histone 1 H2bf and histone 1 H4e, respectively, both of which have a role of binding DNA to form a chromatin structure. Both are replication-dependent histone proteins with expression dependent on cell cycle. Therefore, functional SNPs in this locus might affect the stability of the chromatin structure, varying the cell cycle, cell amount or reaction to inflammation by changing the expression level of histones in the kidney and/or intestine. Since it is also possible that rs11758351 is a surrogate marker near these histone genes, further studies concerning this locus will be necessary.

In this study, we first performed GWASs of gout subtypes, that is, RUE gout and ROL gout (figure 1). From the results of meta-analysis for GWASs of both ROL gout and RUE gout, four shared loci of *GCKR*, *SLC2A9*, *ABCG2* and *CUX2* were identified at a genome-wide significance level, showing the importance of these loci for the pathogenesis of both gout subtypes. Especially for RUE gout, three more loci, *SLC22A12*, *NIPAL1* and *FAM35A*, were identified to be associated at a genome-wide significance level. As described above, it is compatible for *SLC22A12* to be associated with RUE gout, because *SLC22A12* (like *SLC2A9*) encodes a renal urate reabsorption transporter.[25 28]

Of note, *NIPAL1* and *FAM35A* were identified as novel loci by performing GWAS of the RUE gout subtype. Associations with gout and SUA have never been previously reported with *NIPAL1* and *FAM35A*. Furthermore, to our knowledge, there is no study reporting an association between any diseases and *NIPAL1* or *FAM35A*.

NIPAL1, also known as NIPA3, is reportedly expressed on the membrane of some organs including kidney, and to be a magnesium transporter,[23] as another magnesium transporter NIPA2.[23] Because NIPAL1 was associated with RUE gout (ie, gout with renal urate underexcretion), we hypothesised that NIPAL1 is a urate transporter in the human kidney. However, our functional study did not show urate transport via NIPAL1, regardless of the presence of magnesium (figure 3B). Moreover, localisation to the membrane was not detected for NIPAL1 protein, which was mainly expressed within the distal tubules of human kidney, as revealed by immunohistochemical analysis (figure 4A). A similar result was obtained in confocal microscopic observation, with NIPAL1-expressing oocytes showing intracellular localisation of NIPAL1 protein (see online supplementary figure S6). These findings suggest that NIPAL1 is not a urate transporter and that it might be involved in the indirect regulation of urate transport kinetics. Nevertheless, recent studies have revealed associations between hyperuricaemia and magnesium intake,[29] serum magnesium level[30] and magnesium excretion.[31] Together with previous reports, our findings support the hypothesis that there could be some relationship between gout and magnesium handling via magnesium transporters including NIPAL1, and that the present study could well provide new insights into the genetic background of urate and magnesium handling in patients with gout/hyperuricaemia.

FAM35A is ubiquitously expressed in organs including the kidney, and our immunohistochemical analysis of human kidney also revealed cytosolic immunoreactivity of the FAM35A protein mainly in the distal tubules (figure 4B). Our findings from FAM35A and NIPAL1 suggest the involvement of the distal nephron in gout progression as well as dysfunction in urate handling in humans (see online supplementary figure S9). To date, the molecular function of FAM35A is totally unknown. Although further studies are necessary to confirm this, it is possible that genes near *FAM35A* including *GLUD1* (figure 2E)

have some relationship with gout (see also online supplementary note for details).

In addition to studying the Japanese population, we performed a replication study with male Caucasian and NZ Polynesian sample sets for the five newly discovered loci. Since they were not divided into subtypes, further evaluations by meta-analysis were conducted with all gout groups. While other loci were not replicated, rs7903456 of *FAM35A* was replicated with a significant association with gout (table 2). *CUX2*, which was reported by both our present and previous gout GWAS in Japanese,[10] was also replicated in these sample sets (see online supplementary table S5).

A meta-analysis of all gout with Japanese, Caucasian and NZ Polynesian populations for these five SNPs revealed *FAM35A* to be associated with all gout at the genome-wide significance level (figure 5B), and that rs2285340 of *SLC22A12* and rs1165196 of *SLC17A1* showed a significant association but did not reach a genome-wide significance level (see online supplementary figure S7). rs11758351 of *HIST1H2BF-HIST1H4E* and rs11733284 of *NIPAL1* were not associated by this meta-analysis, although these loci showed a genome-wide significant association in the Japanese population. Since this might be due to the differences in LD structure among these populations, a replication analysis with East Asian populations will be necessary for these loci. rs2285340 of *SLC22A12* was monomorphic (only G allele) in Caucasians and not associated with NZ Polynesians. Therefore, replication studies of this locus in East Asian populations would also be insightful for future analysis. Although the underlying molecular mechanism of gout by FAM35A is unknown, this locus seems to have a common pathophysiological risk of gout for Japanese, NZ Polynesians and Caucasians.

In summary, we performed GWASs of all gout as well as gout subtypes and identified five loci in addition to the five loci that we reported previously.[10] Furthermore, the *FAM35A* locus showed an association with all gout by meta-analysis among the Japanese, Caucasian and NZ Polynesian sample sets at a genome-wide level of significance. Together with their expression in the renal distal tubules, the identification of *NIPAL1* and *FAM35A* as gout loci suggests the involvement of the distal nephron (see online supplementary figure S9) in the urate handling of the human kidney and in the pathogenesis of gout/hyperuricaemia. These findings could well provide a clue leading to a novel concept for the therapeutic target of gout (see online supplementary figure S10).

Author affiliations
[1]Department of Integrative Physiology and Bio-Nano Medicine, National Defense Medical College, Tokorozawa, Saitama, Japan
[2]Division of Human Genetics, Department of Integrated Genetics, National Institute of Genetics, Mishima, Shizuoka, Japan
[3]Department of Medical Chemistry, Kurume University School of Medicine, Kurume, Fukuoka, Japan
[4]Department of Dermatology, National Defense Medical College, Tokorozawa, Saitama, Japan
[5]Department of Biochemisty, University of Otago, Dunedin, New Zealand
[6]Department of Pharmacy, The University of Tokyo Hospital, Tokyo, Japan
[7]Department of Human Genetics and Disease Diversity, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan
[8]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Kanagawa, Japan
[9]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka, Japan
[10]Laboratory for Mathematics, National Defense Medical College, Tokorozawa, Saitama, Japan
[11]Department of Biopharmaceutics, School of Pharmacy, Tokyo University of Pharmacy and Life Sciences, Tokyo, Japan
[12]Department of Biopharmaceutics, Graduate School of Pharmaceutical Sciences, Nagoya City University, Nagoya, Aichi, Japan
[13]Department of Preventive Medicine and Public Health, National Defense Medical College, Tokorozawa, Saitama, Japan
[14]Division of Transcriptomics, Research Center for Transomics Medicine, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan
[15]Department of Integrative Genomics, Tohoku Medical Megabank Organization, Tohoku University, Sendai, Miyagi, Japan
[16]Ryougoku East Gate Clinic, Tokyo, Japan
[17]Department of Internal Medicine, Self-Defense Forces Central Hospital, Tokyo, Japan
[18]Program in Radiological and Medical Laboratory Sciences, Pathophysiological Laboratory Sciences, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan
[19]Department of Preventive Medicine, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan
[20]First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Institute of Inherited Metabolic Disorders, Prague, Czech Republic
[21]Institute of Rheumatology, Prague, Czech Republic
[22]Department of Medicine, University of Otago, Christchurch, New Zealand
[23]Department of Medicine, University of Auckland, Grafton, Auckland, New Zealand
[24]Department of Human Physiology and Pathology, Faculty of Pharma-Sciences, Teikyo University, Tokyo, Japan
[25]Department of Internal Medicine, Teikyo University School of Medicine, Tokyo, Japan
[26]Division of Kidney and Hypertension, Department of Internal Medicine, Jikei University School of Medicine, Tokyo, Japan
[27]Department of Pathophysiology and Therapy in Chronic Kidney Disease, Jikei University School of Medicine, Tokyo, Japan
[28]Omics Research Center, National Cerebral and Cardiovascular Center, Suita, Osaka, Japan
[29]Laboratory for Genotyping Development, Center for Integrative Medical Sciences, RIKEN, Yokohama, Kanagawa, Japan
[30]Midorigaoka Hospital, Takatsuki, Osaka, Japan
[31]Kyoto Industrial Health Association, Kyoto, Japan
[32]Department of Pathophysiology, Tokyo University of Pharmacy and Life Sciences, Tokyo, Japan

Acknowledgements The authors thank all the participants involved in this study. They are also grateful to members of the BioBank Japan Project for supporting the study. They are indebted to M Watanabe and Y Katsurada (National Defense Medical College) for immunohistochemical analysis; K Gotanda, Y Morimoto, M Miyazawa, T Chiba, Y Utsumi, S Terashige, Y Kato, H Sasaki, Y Takashima, S Tatsukawa, A Akashi, Y Tanahashi, Y Nagao, M Nakajima, H Inoue, S Takeuchi (National Defense Medical College), M Sonoda (Kurume University School of Medicine) and T Tamatsukuri (Jikei University School of Medicine) for genetic analysis; S Ushida (Ikagaku) and H Fujiwara (Midorigaoka Hospital) for Japanese sample collection; R Akuhata, N Aupouri (Ngati Porou Hauora Charitable Trust) and J H Hindmarsh (Research Coordinator, Ngati Porou Hauora Charitable Trust) for NZ Māori (Eastern Polynesian) sample collection from the Rohe (area) of Ngati Porou iwi; Y Oka, S Kanda and C Umatani (the University of Tokyo) for their biomaterial support and technical advice in the oocyte experiment; J Boocock (University of Otago) for eQTL analysis; H Motohashi (Tohoku University), N Hamajima, M Naito (Nagoya University) and H Tanaka (Aichi Cancer Center Research Institute) for helpful discussion.

Collaborators Members of the Eurogout Consortium are: Mariano Andres (Sección de Reumatología, Hospital General Universitario de Alicante, Alicante), Leo A Joosten (Department of Internal Medicine and Radboud Institute of Molecular Life Science, Radboud University Medical Center, The Netherlands), Matthijs Janssen (Department of Rheumatology, Rijnstate Hospital, The Netherlands), Tim L Jansen (Department of IQ HealthCare, VieCuri Medical Centre, The Netherlands), Frederic Lioté (INSERM, UMR-S 1132, Hospital Lariboisière, Paris, University Paris Diderot (UFR de Médecine), Sorbonne Paris Cité, Paris), Timothy R Radstake (Department of Rheumatology and Clinical Immunology, Laboratory of Translational Immunology, University Medical Centre Utrecht, The Netherlands, and Department of Immunology, University Medical Centre Utrecht, The Netherlands), Philip L Riches (Rheumatic Diseases Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh), Alexander So (DAL, Service of Rheumatology, Laboratory of Rheumatology, University of Lausanne, CHUV, Nestlé), Anne-Kathrin Tausche (Department of Rheumatology, University Clinic 'Carl-Gustav-Carus', Dresden).

Contributors AN, HNakaoka, KY, MS, AS, YT, TT, NS, TRM and HM conceived and designed this study. YSakurai, HS, II, ATakahashi and MKubo assisted with research design. AN, HNakaoka, MS, YOkada, YKamatani, THigashino, YKawamura, ATokumasu, KO, TK, KW, BS, KP, ATakahashi, MKubo, HOoyama, TS, KIchida and HM collected and analyzed clinical data of Japanese participants. AS, LKS, ND, Eurogout Consortium and TRM collected and analyzed clinical data of replication participants. AN, KY, MS, AS, YShirahama, SS, THigashino, YKawamura, HOgata, MKawaguchi, ID, NS, TRM and HM performed genetic analysis. AN, HNakaoka, MS,

AS, YOkada, YKamatani, TN, HNakashima, ATakahashi, TRM and HM performed statistical analysis. AN, YT, TT, KInoue, TYasujima, HY, HS and HM performed functional analysis and localization analysis. AN, HNakaoka, KY, MS, AS, YT, YOkada, YKamatani, TN, TT, KInoue, TYasujima, HY, YOhkawa, NS, TRM and HM analyzed data. ID, TI, MH, SF, TYokoo, THosoya, KIchida provided intellectual input and assisted with the preparation of the manuscript. AN, HNakaoka, KY, MS, AS, YT, NS, TRM and HM wrote the manuscript.

Funding This study was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, including MEXT KAKENHI (Nos. 25293145 and 15K15227), Grants-in-Aid for Scientific Research on Priority Areas (No. 17015018) and Innovative Areas (Nos. 221S0001 and 221S0002) and a JSPS KAKENHI Grant (Nos. 16H06277 and 16H06279), the Ministry of Health, Labour and Welfare of Japan, the Ministry of Defense of Japan, the Japan Society for the Promotion of Science, the Kawano Masanori Memorial Foundation for Promotion of Pediatrics, the Gout Research Foundation of Japan and the Health Research Council of New Zealand. The BioBank Japan Project was supported by MEXT of Japan.

Competing interests TT, KIchida, NS and HM have a patent pending based on the work reported in this paper.

Patient consent Obtained.

Ethics approval This study was approved by the institutional ethical committees, and written consent was obtained from all of its participants. All involved procedures were performed in accordance with the Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Smith E, Hoy D, Cross M, *et al*. The global burden of gout: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis* 2014;73:1470–6.
2 Ichida K, Matsuo H, Takada T, *et al*. Decreased extra-renal urate excretion is a common cause of hyperuricemia. *Nat Commun* 2012;3:764.
3 Wortmann RL. Disorders of purine and pyrimidine metabolism. In: Fauci AS, Braunwald E, Kasper D, *et al*., eds. *Harrison's principles of internal medicine*. 17th edn. New York, NY: McGraw-Hill, 2008:2444–9.
4 Matsuo H, Nakayama A, Sakiyama M, *et al*. ABCG2 dysfunction causes hyperuricemia due to both renal urate underexcretion and renal urate overload. *Sci Rep* 2014;4:3755.
5 Woodward OM, Köttgen A, Coresh J, *et al*. Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proc Natl Acad Sci USA* 2009;106:10338–42.
6 Matsuo H, Takada T, Ichida K, *et al*. Common defects of ABCG2, a high-capacity urate exporter, cause gout: a function-based genetic analysis in a Japanese population. *Sci Transl Med* 2009;1:5ra11.
7 Matsuo H, Ichida K, Takada T, *et al*. Common dysfunctional variants in *ABCG2* are a major cause of early-onset gout. *Sci Rep* 2013;3:2014.
8 Nakayama A, Matsuo H, Nakaoka H, *et al*. Common dysfunctional variants of ABCG2 have stronger impact on hyperuricemia progression than typical environmental risk factors. *Sci Rep* 2014;4:5227.
9 Phipps-Green AJ, Merriman ME, Topless R, *et al*. Twenty-eight loci that influence serum urate levels: analysis of association with gout. *Ann Rheum Dis* 2016;75:124–30.
10 Matsuo H, Yamamoto K, Nakaoka H, *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.
11 Nakamura T, Shi D, Tzetis M, *et al*. Meta-analysis of association between the ASPN D-repeat and osteoarthritis. *Hum Mol Genet* 2007;16:1676–81.
12 Kamatani Y, Matsuda K, Okada Y, *et al*. Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nat Genet* 2010;42:210–15.
13 Rasheed H, Phipps-Green A, Topless R, *et al*. Association of the lipoprotein receptor-related protein 2 gene with gout and non-additive interaction with alcohol consumption. *Arthritis Res Ther* 2013;15:R177.
14 Lioté F, Merriman T, Nasi S, *et al*. Workshop report: 4th European crystal network meeting. *Arthritis Res Ther* 2013;15:304.
15 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
16 Chiba T, Matsuo H, Kawamura Y, *et al*. NPT1/SLC17A1 is a renal urate exporter in humans and its common gain-of-function variant decreases the risk of renal underexcretion gout. *Control Clin Trials* 2015;67:281–7.
17 Matsuo H, Kanai Y, Kim JY, *et al*. Identification of a novel Na^+-independent acidic amino acid transporter with structural similarity to the member of a heterodimeric amino acid transporter family associated with unknown heavy chains. *J Biol Chem* 2002;277:21017–26.
18 Matsuo H, Tsukada S, Nakata T, *et al*. Expression of a system L neutral amino acid transporter at the blood-brain barrier. *Neuroreport* 2000;11:3507–11.
19 Matsuo H, Kanai Y, Tokunaga M, *et al*. High affinity D- and L-serine transporter Asc-1: cloning and dendritic localization in the rat cerebral and cerebellar cortices. *Neurosci Lett* 2004;358:123–6.
20 Hollis-Moffatt JE, Phipps-Green AJ, Chapman B, *et al*. The renal urate transporter SLC17A1 locus: confirmation of association with gout. *Arthritis Res Ther* 2012;14:R92.
21 Diodato D, Melchionda L, Haack TB, *et al*. VARS2 and TARS2 mutations in patients with mitochondrial encephalomyopathies. *Hum Mutat* 2014;35:983–9.
22 Köttgen A, Albrecht E, Teumer A, *et al*. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat Genet* 2013;45:145–54.
23 Goytain A, Hines RM, Quamme GA. Functional characterization of NIPA2, a selective Mg^{2+} transporter. *Am J Physiol, Cell Physiol* 2008;295:C944–53.
24 Li C, Li Z, Liu S, *et al*. Genome-wide association analysis identifies three new risk loci for gout arthritis in Han Chinese. *Nat Commun* 2015;6:7041.
25 Enomoto A, Kimura H, Chairoungdua A, *et al*. Molecular identification of a renal urate anion exchanger that regulates blood urate levels. *Nature* 2002;417:447–52.
26 Kolz M, Johnson T, Sanna S, *et al*. Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS Genet* 2009;5:e1000504.
27 Okada Y, Sim X, Go MJ, *et al*. Meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations. *Nat Genet* 2012;44:904–9.
28 Matsuo H, Chiba T, Nagamori S, *et al*. Mutations in glucose transporter 9 gene *SLC2A9* cause renal hypouricemia. *Am J Hum Genet* 2008;83:744–51.
29 Wang YL, Zeng C, Wei J, *et al*. Association between dietary magnesium intake and hyperuricemia. *PLoS ONE* 2015;10:e0141079.
30 Zeng C, Wang YL, Wei J, *et al*. Association between low serum magnesium concentration and hyperuricemia. *Magnes Res* 2015;28:56–63.
31 Kuroczycka-Saniutycz E, Porowski T, Protas PT, *et al*. Does obesity or hyperuricemia influence lithogenic risk profile in children with urolithiasis? *Pediatr Nephrol* 2015;30:797–803.

CONCISE REPORT

An induction or flare of arthritis and/or sacroiliitis by vedolizumab in inflammatory bowel disease: a case series

G Varkas,[1,2] K Thevissen,[1,3] G De Brabanter,[4] L Van Praet,[1] F Czul-gurdian,[5] H Cypers,[1,2] J De Kock,[1] P Carron,[1,3] M De Vos,[6] P Hindryckx,[6] J Arts,[7] I Vanneuville,[3] P Schoenaers,[8] B Claerhout,[8] M Abreu,[5] F Van den Bosch,[1] D Elewaut[1,2]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210233).

For numbered affiliations see end of article.

Correspondence to
Professor Dr D Elewaut, Department of Rheumatology, Ghent University Hospital, De Pintelaan 185, Ghent 9000, Belgium; Dirk.elewaut@ugent.be

Received 18 July 2016
Revised 25 October 2016
Accepted 5 November 2016
Published Online First 29 November 2016

ABSTRACT

Background In inflammatory bowel disease (IBD), a new biological therapy has recently been approved. Vedolizumab is a humanised IgG1 monoclonal antibody to α4β7 integrin that modulates gut lymphocyte trafficking. Although an exclusively local effect of vedolizumab could be expected based on the restricted presence of the α4β7–mucosal vascular addressin cell adhesion molecule 1 complex in the gut, past combined success with anti-tumour necrosis factor, and previous demonstration of α4β7 integrin in the joint, led to the expectation of a therapeutic efficacy in spondyloarthritis. Nonetheless, the effect of vedolizumab on extraintestinal manifestations—and especially the joint—has not been reported so far.

Case report A series of five patients with IBD who were treated with vedolizumab and promptly developed new onset or exacerbation of sacroiliitis or arthritis are reported.

Conclusions Vedolizumab therapy does not seem to show any efficacy in and might even induce arthritis and/or sacroiliitis. However, larger cohort studies are needed to provide information on the prevalence, the evolution and underlying mechanism.

INTRODUCTION

Spondyloarthritis (SpA) and inflammatory bowel disease (IBD) have clinical, imaging and genetic evidence supporting overlap in their pathogenesis, which is best reflected in the shared efficacy of anti-tumour necrosis factor (TNF). In IBD, a new humanised IgG1 monoclonal antibody to α4β7 has been approved, which modulates gut lymphocyte trafficking. In clinical trials, vedolizumab induced a clinical response in Crohn's disease (CD) and ulcerative colitis (UC) in respectively one-third and up to half of treated patients.[1 2] Although an exclusively local effect could be expected based on the restricted presence of the α4β7–mucosal vascular addressin cell adhesion molecule 1 (MadCAM) complex in the gut,[3 4] combined success with anti-TNF, and previous demonstration of α4β7 in the joint,[5–9] led to an anticipated efficacy in SpA. Nonetheless, the effect on extraintestinal manifestations—and especially the joint—has not been reported.

MATERIALS AND METHODS

Information on the recruitment of patients and requirements for inclusion can be found in the online supplementary file.

Case 1

A 50-year-old woman had been diagnosed with CD and psoriasis at the age of 39. Two months after initiation of vedolizumab, she reported progressive inflammatory back pain with 30 min morning stiffness, which responded well to non-steroidal anti-inflammatory drugs (NSAIDs). MRI of the sacroiliac (SI) joints showed severe bilateral sacroiliitis. At follow-up after 6 months, the patient showed decrease of Bone Marrow Edema (BME) with etoricoxib 90 mg daily (figure 1).

Case 2

A 28-year-old woman diagnosed with UC at the age of 16 was referred for intermittent pain in the lower limbs after treatment with vedolizumab. She had not been diagnosed with SpA, nor did she present with other SpA features. The patient presented with a painful left shoulder and arthritis of the right wrist. On ultrasound, intercarpal effusion, synovial hyperproliferation and power Doppler (PD) of 1–2 was present, in addition to severe PD signal in the tendon sheath of the flexor pollicis longus and flexor carpi radialis muscle.

Case 3

A 30-year-old man had been diagnosed with ankylosing spondylitis (AS) and CD at the age of 25. He had been treated with golimumab and infliximab until august 2015 and was in clinical remission. Due to secondary inefficacy at the level of the gut, the patient was switched to vedolizumab. Four weeks later, he presented with arthralgia, back pain, pain at night and 90 min morning stiffness with elevated C reactive protein. Furthermore, MRI unveiled acute inflammatory lesions of the axial skeleton (figure 2).

Case 4

A 47-year-old woman, diagnosed with CD at the age of 42, consulted for intense low back pain following treatment with vedolizumab. The pain was situated at the right side, worse in the early

To cite: Varkas G, Thevissen K, De Brabanter G, *et al*. *Ann Rheum Dis* 2017;**76**:878–881.

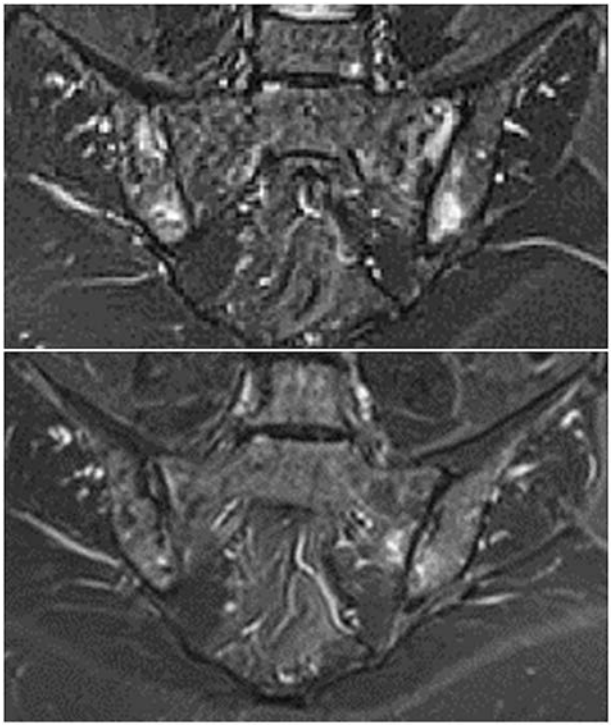


Figure 1 MRI of the sacroiliac joints (SIJs) in a 50-year-old patient with Crohn's disease presenting with inflammatory back pain after vedolizumab treatment. Top: bilateral sacroiliitis at initial presentation. Bottom: decrease of BME of the SIJs after 6 months of etoricoxib 90 mg daily.

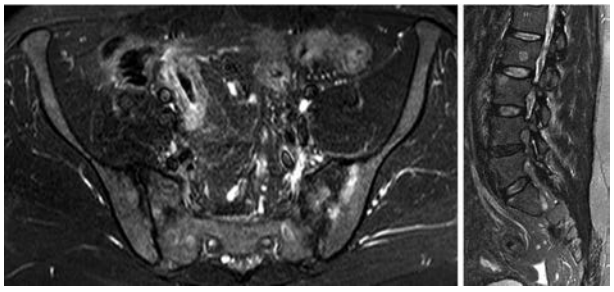


Figure 2 MRI of the sacroiliac joints (SIJs) and spine in a 30-year-old patient with AS and Crohn's disease presenting with increased inflammatory back pain after vedolizumab treatment. Left: BME of the left SIJ. Right: inflammatory corner lesions of the spine at level S1, L4 and L3.

morning and resulted in awakening at night. She displayed no other features of SpA, nor had she been diagnosed with SpA in the past. MRI of the SI joints displayed unilateral sacroiliitis of the right SI joint.

Case 5

A 26-year-old woman was diagnosed with UC at the age of 18. She presented with polyarticular joint pain after initiation of vedolizumab, predominantly in the lower limbs with moderate response to low-dose NSAIDs. At clinical examination, a painful, warm left wrist and right elbow, and synovitis of several metatarsophalangeal joints of the right foot were detected, in combination with tenderness at the Achilles tendon enthesis of the left heel. Blood analysis confirmed an inflammatory aetiology. Treatment strategies and response to treatment for all cases are shown in table 1.

DISCUSSION

This is the first report on new onset or exacerbation of arthritis/sacroiliitis in vedolizumab-treated patients. Despite the proven efficacy in IBD, both in anti-TNF-naive[1 2] and in anti-TNF-exposed patients,[10] α4β7 blockade seems to facilitate synovitis with similar distribution to SpA in some patients, irrespective of the response to treatment at the level of the gut. According to the European Medicines Agency, arthralgia, back pain and pain in the extremities were reported common (≥1/100) to very common (≥1/10) in clinical trials. Yet, no mention of proven arthritis or sacroiliitis was made. One can speculate that these may have been present, but were insufficiently investigated. Surprisingly, in trials with natalizumab, a non-selective inhibitor of α4, solely arthralgia was reported. Meanwhile, a large safety study could not demonstrate higher rates of arthralgia in patients treated with vedolizumab compared with placebo.[11] Similarly, no mention was made of increased musculoskeletal symptoms with natalizumab in the ENCORE trial, nor in multiple sclerosis.[12 13]

Lymphocyte trafficking is facilitated by adhesion molecules, divided into integrins, selectins and the immunoglobulin superfamily. The integrins consist of an α-subunit and β-subunit, of which α4 forms a heterodimer with β1 or β7. The α4β1 serves as a ligand for vascular cell adhesion molecule-1 (VCAM-1) and is expressed on leucocytes and endothelial cells, whereas α4β7 integrin serves as a ligand for both MadCAM-1 and VCAM-1, and is expressed on a subset of lymphocytes such as CD4+ and CD8+ T cells. Indirect evidence in humans has shown that T cells with high α4β7 expression have a predilection for the gut mucosa. Similarly, MadCAM-1 is selectively expressed on the mucosal lymphoid organ high endothelial venules (HEVs) and on the gut lamina propria venules.[4 6 14] In particular, vedolizumab has been shown to inhibit the interaction between α4β7 and MadCAM-1, but not VCAM-1.[15]

One of the many hypotheses is that integrins and adhesion molecules play a role in the interception of recirculating activated lymphocytes between the gut and the synovial membrane due to the inhibition of the α4β7 integrin homing at the level of the gut. As the preferred interaction could not take place, these activated cells could easily drift across tissues in search of a landing as the overall survival of cells was not affected.[16] Even though both axial and peripheral disease manifested in our patients, irrefutable data on the presence of such adhesion molecules at the level of the spine are lacking. Nevertheless, MadCAM-1 was found to be upregulated in the HEVs of bone marrow in a small sample of patients with active axial SpA.[14] Although MadCAM-1 generally appears to be restricted to the gut, α4β7 has been demonstrated in the inflamed joint with increased expression in various inflammatory conditions compared with healthy controls as a consequence of exposure to inflammatory cytokines.[5–9 17] Alternatively, in the presence of vedolizumab, cellular recruitment may be mediated by yet to be determined adhesion molecules. This recirculation theory might explain the short mean interval of 64 days between vedolizumab initiation and the expression of symptoms. The remarkable discrepancy in joint symptoms by natalizumab compared with vedolizumab, and the human leukocyte antigen (HLA)-B27 negativity of these cases, may point towards a predominantly innate immune mechanism of joint disease. In IBD, the prevalence of HLA-B27 is comparable to healthy controls. However, in association with SpA, the prevalence is much

Table 1 Patient characteristics, spondyloarthritis (SpA) features, treatment and outcome in vedolizumab-treated patients with inflammatory bowel disease (IBD) presenting with arthritis or sacroiliitis

	Gender	Age	HLA-B27	Smoker	Concomitant medication	IBD	Response of gut inflammation	Prior diagnosis of SpA	TTF (days)	SpA feature	Therapy	Outcome of SpA feature at 6 months
1	F	50	Neg	No	Mesalazine	CD	Good HBI 0 at week 14	No	60	Sacroiliitis	Single IA SIJ, chronic (Cox2 selective) NSAID	Clinical remission and imaging improvement
2	F	28	Neg	No	AZA	UC	Good Mayo 0 at week 10	No	58	Arthritis	Single IA	Clinical and imaging remission
3	M	30	Neg	No	AZA	CD	Good HBI 0 at week 28	Yes (axial disease)	14	Sacroiliitis	Chronic (Cox2 selective) NSAID	Active disease
4	F	47	Neg	Yes	Mesalazine	CD	Good HBI 2 at week 20+HBI 1 at week 32	No	114	Sacroiliitis	Single IA SIJ	Clinical remission
5	F	26	Neg	No	None	UC	Poor Mayo 3 at week 10	No	73	Arthritis	OCS	Clinical remission

AZA, azathioprine 100 mg; CD, Crohn's disease; F, female; HBI, Harvey-Bradshaw Index; IA, intra-articular infiltration; M, male; NSAID, non-steroidal anti-inflammatory drugs; OCS, oral corticosteroids; SIJ, sacroiliac joints; TTF, time to flare; UC, ulcerative colitis.

higher.[18] Notwithstanding, these interactions may be volatile and/or display low affinity, or be subject to a distinct pattern of interactions, which might explain why these features do not manifest in every patient.

Nevertheless, we cannot rule out that vedolizumab simply does not have any efficacy in SpA, which can be explained by vedolizumab solely intervening in the interaction of α4β7 with gut-specific MadCAM-1.[6] The overlap between SpA and IBD manifests in up to 30% of patients with IBD, and in patients with concomitant disease, gut flares have been associated with joint flares.[19] Admittedly, these joint flares may have occurred before any efficacy by vedolizumab at the level of the gut had taken place. However, the effect of vedolizumab on the gut did not seem to be related to the joint symptoms as a good response at the level of the gut did not necessarily result in a better outcome of the SpA feature over time, which demonstrates at least some level of disconnect. Although vedolizumab therapy was not discontinued in these patients due to a lack of alternative therapeutic options regarding their IBD, more than half of these patients necessitated additional chronic therapy in order to reach acceptable SpA disease activity.

Moreover, due to the prior exposure to anti-TNF, underlying SpA might have been suppressed in the past. This inefficacy theory is supported by the flare of sacroiliitis in the patient with pre-existing SpA in remission under anti-TNF therapy. Still, in the remaining four patients, the time frame in which symptoms occur, the dissociation of gut and joint response, and the absence of HLA-B27, family history or other SpA features render this lack of efficacy theory less likely as opposed to the induction of SpA features.

In conclusion, vedolizumab does not seem to show any efficacy in and might even induce arthritis and/or sacroiliitis. Larger cohort studies are needed to provide information on the prevalence, evolution and underlying mechanism.

Author affiliations
[1]Department of Rheumatology, Ghent University Hospital, Ghent, Belgium
[2]VIB Inflammation Research Center, Ghent University, Ghent, Belgium
[3]Department of Rheumatology/Physical medicine and rehabilitation, AZ Alma, Sijsele-Eeklo, Belgium
[4]Department of Rheumatology, AZ Sint-Lucas Brugge, Bruges, Belgium
[5]The Crohn's and Colitis Center, Miller School of Medicine, University of Miami, Miami, Florida, USA
[6]Department of Gastroenterology, Ghent University Hospital, Ghent, Belgium
[7]Department of Gastroenterology, AZ Sint-Lucas Brugge, Bruges, Belgium
[8]Department of Gastroenterology, AZ Alma, Sijsele-Eeklo, Belgium

Acknowledgements The investigators would like to thank their patients for giving consent.

Contributors GV, MDV, PH, FVdB and DE: study concept and design. GV, KT, GDB, LVP, FC, HC, JDK, PC, MDV, PH, JA, IV, PS, BC, MA, FVdB and DE: data acquisition. GV, PC, FVdB and DE: analysis and interpretation of data. GV, MDV, MA, FVdB and DE: manuscript preparation. GV, KT, GDB, LVP, MDV, PH, MA, FVdB and DE: manuscript revision.

Funding This study was supported by a grant of Ghent University to GV. DE is supported by a fund of Scientific Research–Flanders (FWO) and the Research Council of Ghent University. DE is also a member of a multidisciplinary research platform (MRP) of Ghent University and is supported by Interuniversity Attraction Pole (IUAP) grant Devrepair from the Belspo Agency (project P7/07).

Competing interests DE has received grants or speakers fees from Boehringer Ingelheim, Pfizer, UCB, Merck, Novartis, Janssen and Abbvie. FVdB received consultancy and/or speaker fees from Abbvie, Celgene, Janssen, Merck, Novartis, Pfizer and UCB.

Patient consent Obtained.

Ethics approval The study protocol was approved by the Ethical Committee of Ghent University Hospital.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Feagan BG, Rutgeerts P, Sands BE, *et al*. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2013;369:699–710.
2 Sandborn WJ, Feagan BG, Rutgeerts P, *et al*. Vedolizumab as induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2013;369:711–21.
3 Wyant T, Leach T, Sankoh S, *et al*. Vedolizumab affects antibody responses to immunisation selectively in the gastrointestinal tract: randomised controlled trial results. *Gut* 2015;64:77–83.
4 Schweighoffer T, Tanaka Y, Tidswell M, *et al*. Selective expression of integrin alpha 4 beta 7 on a subset of human CD4+ memory T cells with Hallmarks of gut-trophism. *J Immunol* 1993;151:717–29.
5 Salmi M, Andrew DP, Butcher EC, *et al*. Dual binding capacity of mucosal immunoblasts to mucosal and synovial endothelium in humans: dissection of the molecular mechanisms. *J Exp Med* 1995;181:137–49.
6 Salmi M, Rajala P, Jalkanen S. Homing of mucosal leukocytes to joints. Distinct endothelial ligands in synovium mediate leukocyte-subtype specific adhesion. *J Clin Invest* 1997;99:2165–72.
7 Elewaut D, De Keyser F, Van Den Bosch F, *et al*. Enrichment of T cells carrying beta7 integrins in inflamed synovial tissue from patients with early spondyloarthropathy, compared to rheumatoid arthritis. *J Rheumatol* 1998;25:1932–7.
8 Salmi M, Jalkanen S. Endothelial ligands and homing of mucosal leukocytes in extraintestinal manifestations of IBD. *Inflamm Bowel Dis* 1998;4:149–56.
9 Salmi M, Jalkanen S. Human leukocyte subpopulations from inflamed gut bind to joint vasculature using distinct sets of adhesion molecules. *J Immunol* 2001;166:4650–7.
10 Sands BE, Feagan BG, Rutgeerts P, *et al*. Effects of vedolizumab induction therapy for patients with Crohn's disease in whom tumor necrosis factor antagonist treatment failed. *Gastroenterology* 2014;147:618–27.e3.
11 Colombel JF, Sands BE, Rutgeerts P, *et al*. The safety of vedolizumab for ulcerative colitis and Crohn's disease. *Gut* Published Online First: 18 Feb 2016. doi: 10.1136/gutjnl-2015-311079
12 Targan SR, Feagan BG, Fedorak RN, *et al*. Natalizumab for the treatment of active Crohn's disease: results of the ENCORE Trial. *Gastroenterology* 2007;132:1672–83.
13 Polman CH, O'Connor PW, Havrdova E, *et al*. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med* 2006;354:899–910.
14 Ciccia F, Guggino G, Rizzo A, *et al*. Type 3 innate lymphoid cells producing IL-17 and IL-22 are expanded in the gut, in the peripheral blood, synovial fluid and bone marrow of patients with ankylosing spondylitis. *Ann Rheum Dis* 2015;74:1739–47.
15 Wyant T, Fedyk E, Abhyankar B. An overview of the mechanism of action of the monoclonal antibody vedolizumab. *J Crohns Colitis* 2016;10:1437–44.
16 Feagan BG, Greenberg GR, Wild G, *et al*. Treatment of ulcerative colitis with a humanized antibody to the alpha4beta7 integrin. *N Engl J Med* 2005;352:2499–507.
17 Morales-Ducret J, Wayner E, Elices MJ, *et al*. Alpha 4/beta 1 integrin (VLA-4) ligands in arthritis. Vascular cell adhesion molecule-1 expression in synovium and on fibroblast-like synoviocytes. *J Immunol* 1992;149:1424–31.
18 de Vlam K, Mielants H, Cuvelier C, *et al*. Spondyloarthropathy is underestimated in inflammatory bowel disease: prevalence and HLA association. *J Rheumatol* 2000;27:2860–5.
19 De Vos M, Mielants H, Cuvelier C, *et al*. Long-term evolution of gut inflammation in patients with spondyloarthropathy. *Gastroenterology* 1996;110:1696–703.

CONCISE REPORT

The risk of fracture among patients with psoriatic arthritis and psoriasis: a population-based study

Alexis Ogdie,[1] Lauren Harter,[2] Daniel Shin,[3] Joshua Baker,[4] Junko Takeshita,[5] Hyon K Choi,[6] Thorvardur Jon Love,[7] Joel M Gelfand[5]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210441).

For numbered affiliations see end of article.

Correspondence to
Dr Alexis Ogdie, Division of Rheumatology, Center for Clinical Epidemiology and Biostatistics, Center for Pharmacoepidemiology Research and Training, Perelman School of Medicine at the University of Pennsylvania, White Building, Room 5023, 3400 Spruce St., Philadelphia, PA 19104, USA; alexis.ogdie@uphs.upenn.edu

AO and LH contributed equally.

Received 30 August 2016
Revised 7 December 2016
Accepted 28 December 2016
Published Online First 16 January 2017

ABSTRACT

Objective To determine the risk of fracture and osteoporosis among patients with psoriatic arthritis (PsA) and psoriasis, compared with the general population and patients with rheumatoid arthritis (RA).

Methods A population-based cohort study was performed in The Health Improvement Network in the UK using data from 1994 to 2014. Patients aged 18–89 years with PsA or psoriasis and up to five unexposed controls matched by practice and start date within that practice were included. Patients with RA and matched controls were included for comparison. Severe psoriasis was defined by a code for psoriasis and either phototherapy or a systemic medication for psoriasis. Incidence and adjusted HRs (aHR) for fracture (all, hip, vertebral) were calculated.

Results Patients with PsA (n=9788), psoriasis (n=158 323) and controls (n=821 834) were identified. Patients with PsA had an elevated risk of all fracture aHR 1.26 (1.06 to 1.27). Patients with mild psoriasis had elevated risk of all fractures, vertebral and hip fracture: aHR 1.07 (1.05 to 1.10), 1.17 (1.03 to 1.33) and 1.13 (1.04 to 1.22). Patients with severe psoriasis had significantly elevated risk of all fracture and vertebral fracture: aHR 1.26 (1.15 to 1.39) and 2.23 (1.54 to 3.22).

Conclusions PsA and psoriasis are associated with an elevated risk for fracture.

INTRODUCTION

Osteoporosis (OP) is one of the most common and costly diseases: one in every two women and one in five men will experience a fracture after the age of 50 years.[1] Hospitalisations for OP fracture are more common than hospitalisations for myocardial infarction and stroke combined[2] and fractures result in pain, immobility, nursing home placement, isolation and depression, in addition to other health problems.[1] Rheumatoid arthritis (RA) is a known risk factor for OP.[3] Ankylosing spondylitis (AS), despite its association with new bone formation and syndesmophytes, is also associated with vertebral OP and fractures.

While psoriatic arthritis (PsA) and psoriasis demonstrate a similar Th1-driven and Th17-driven inflammation to RA and a pathophysiological link to AS, few studies have addressed the risk of OP in these patients[4 5] Two studies have reported an increased prevalence of osteopenia or OP among patients with psoriasis in Taiwan and Israel.[6 7] Additional studies examined bone mineral density in patients with PsA compared with healthy controls, though with conflicting results.[8–12] These studies have been limited by cross-sectional designs and lack of adjustment for obesity, smoking or other risk factors for OP. To our knowledge, no studies have evaluated the risk of incident fracture in PsA or psoriasis. Therefore, the objective of this study was to examine the incidence of fracture in patients with PsA and psoriasis and compare this with matched controls from the general population and patients with RA.

METHODS

Study design

We performed a longitudinal cohort study to examine the risk of incident fracture among patients with psoriasis and PsA compared with patients from the general population and patients with RA.

Data source

Data from The Health Improvement Network (THIN) in the UK between 1994 and January 2014 were used.[35]

Study population

All patients with PsA or psoriasis between the ages of 18 and 89 at the start date were included if they had observation time in THIN after Vision software implementation. Patients were excluded if they died or transferred out of the practice prior to the implementation of Vision software. Patients with a history of fracture or OP or a history of bisphosphonate prescriptions were excluded. Patients with psoriasis, PsA and RA were matched to up to five unexposed controls from the general population (matching is described in the online supplementary methods).

Exposure and outcome definitions

PsA, psoriasis and RA were defined by the presence of at least one read code consistent with these diseases using previously validated codes.[13–17] Severe psoriasis was defined as a code for psoriasis plus a code for either phototherapy or a systemic medication for psoriasis. The outcomes of interest were fractures (all fractures, hip fracture and vertebral fracture).[18–21] Disease-modifying antirheumatic drugs (DMARDs) and covariates are listed in the online supplementary methods.

Person-time calculation

Cohort time started at the latest of the following: diagnosis with psoriasis, PsA or RA (diagnosis date

To cite: Ogdie A, Harter L, Shin D, *et al*. *Ann Rheum Dis* 2017;**76**:882–885.

for unexposed controls was the encounter date within 6 months of the matched patient's diagnosis date), 180 days after registration in the practice or Vision date (software implementation in the practice). Cohort time ended at earliest of development of the outcome, transfer out of the practice, practice stops contributing to THIN, death or the end of the study. All covariates of interest were measured prior to cohort entry.

Statistical analysis

Descriptive statistics were used to examine age, sex, person-time and covariate distribution between patients with PsA, psoriasis and unexposed controls. The number of events and cumulative incidence of fracture were calculated for each group. Cox proportional hazards models were used to calculate unadjusted and adjusted HRs with 95% CIs. A purposeful selection modelling approach was used to determine in the most biologically plausible and parsimonious model.[22] The proportional hazards assumption was assessed using log–log plots. Sensitivity analyses are described in the online supplementary data.

Ethics review

This study was approved by the University of Pennsylvania Institutional Review Board and the Cegedim Scientific Review Committee. This paper was prepared according to STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines.[23]

RESULTS

After applying inclusion and exclusion criteria, 9788 patients with PsA, 158 323 patients with psoriasis and 821 834 matched controls were identified. Baseline demographics are shown in table 1 and online supplementary table S1. Time in the cohort was similar among the groups. Approximately 5% of patients with psoriasis had been prescribed a DMARD or received phototherapy. Oral corticosteroids were prescribed for 17% of patients with PsA, 21% of patients with severe 9% of patients with mild psoriasis and controls. Proton pump inhibitors were commonly prescribed for patients with PsA and severe psoriasis (31% each) than controls and those with mild psoriasis (both 15%). Patients with PsA and psoriasis were also more likely to have been prescribed an antidepressant in the baseline period and had a higher prevalence of diabetes than controls. Those with mild or severe psoriasis were more likely to be current smokers, while patients with PsA and severe psoriasis had higher rates of heavy alcohol use.

The number of events, unadjusted incidence per 10 000 person-years and HRs for fracture are presented in table 2. Results for RA are included in the table for comparison. After

Table 1 Baseline demographics

		Control	PsA	Mild psoriasis	Severe psoriasis	RA
	n	821 834	9788	149 809	8514	39 306
Age	Mean (SD)	50.18 (17.47)	49.74 (14.09)	46.67 (17.43)	49.29 (15.18)	58.71 (15.33)
Female sex	n (%)	469 431 (57.12%)	5029 (51.38%)	79 961 (53.38%)	4498 (52.83%)	27 198 (69.20%)
Cohort time*	Mean (SD)	6.75 (4.87)	6.17 (4.67)	6.37 (4.80)	5.50 (4.19)	6.29 (4.67)
Visits in 1 year before start	Mean (SD)	4.86 (6.45)	8.03 (8.81)	6.22 (6.72)	10.54 (10.13)	10.02 (9.61)
Cancer†	n (%)	121 488 (14.78%)	1218 (12.44%)	20 546 (13.71%)	1264 (14.85%)	6092 (15.50%)
Chronic kidney disease	n (%)	17 335 (2.11%)	188 (1.92%)	2426 (1.62%)	249 (2.92%)	1332 (3.39%)
Atrial fibrillation	n (%)	16 923 (2.06%)	160 (1.63%)	2425 (1.62%)	142 (1.67%)	1263 (3.21%)
Diabetes	n (%)	49 554 (6.03%)	743 (7.59%)	8091 (5.40%)	768 (9.02%)	3204 (8.15%)
Cardiovascular disease	n (%)	48 647 (5.92%)	458 (4.68%)	7640 (5.10%)	479 (5.63%)	3712 (9.44%)
COPD	n (%)	17 585 (2.14%)	207 (2.11%)	3262 (2.18%)	233 (2.74%)	1654 (4.21%)
Liver disease	n (%)	9653 (1.17%)	182 (1.86%)	1947 (1.30%)	163 (1.91%)	594 (1.51%)
Dementia	n (%)	3963 (0.48%)	32 (0.33%)	788 (0.53%)	30 (0.35%)	357 (0.91%)
Stroke	n (%)	21 827 (2.66%)	205 (2.09%)	3450 (2.30%)	212 (2.49%)	1613 (4.10%)
Antidepressant use	n (%)	178 630 (21.74%)	2861 (29.23%)	33 252 (22.20%)	2834 (33.29%)	11 805 (30.03%)
Antiepileptic use	n (%)	25 338 (3.08%)	398 (4.07%)	4774 (3.19%)	424 (4.98%)	1877 (4.78%)
Oral corticosteroids	n (%)	77 521 (9.43%)	1645 (16.81%)	12 811 (8.55%)	1819 (21.36%)	11 532 (29.34%)
PPI use	n (%)	125 493 (15.27%)	3021 (30.86%)	22 615 (15.10%)	2666 (31.31%)	13 408 (34.11%)
Hormone therapy‡	n (%)	252 829 (30.76%)	2643 (27.00%)	40 218 (26.85%)	2521 (29.61%)	11 158 (28.39%)
Smoking	Never/former (n, %)	561 085 (68.27%)	6966 (71.17%)	92 044 (61.44%)	5658 (66.46%)	26 497 (67.41%)
	Current (n, %)	170 562 (20.75%)	2011 (20.55%)	41 751 (27.87%)	2322 (27.27%)	8676 (22.07%)
	Missing (n, %)	*90 187 (10.97%)*	*811 (8.29%)*	*16 041 (10.69%)*	*534 (6.27%)*	*4133 (10.51%)*
Alcohol use	Never (n, %)	96 244 (11.72%)	1150 (11.75%)	16 023 (10.70%)	973 (11.43%)	6503 (16.54%)
	Some (n, %)	519 079 (63.16%)	6462 (66.02%)	97 034 (64.77%)	5600 (65.77%)	23 137 (58.86%)
	Heavy (n, %)	31 735 (3.86%)	477 (4.87%)	5501 (3.67%)	552 (6.48%)	2167 (5.51%)
	Missing (n, %)	*174 676 (21.25%)*	*1699 (17.36%)*	*31 251 (20.86%)*	*1389 (16.31%)*	*7499 (19.08%)*
BMI	*Mean (SD)*	26.39 (5.46)	28.03 (5.86)	26.65 (5.60)	28.08 (6.11)	26.69 (5.55)
	Missing (n, %)	*168 709 (20.53%)*	*1659 (16.95%)*	*30 833 (20.58%)*	*1327 (15.59%)*	*7446 (18.94%)*

*Time from index date to end date.
†Cancer includes haematologic malignancy and solid tumour malignancies.
‡Hormone therapy refers to the use of oral contraceptives as well as hormone replacement therapy.
BMI, body mass index; COPD, chronic obstructive pulmonary disease; PPI, proton pump inhibitor; PsA, psoriatic arthritis; RA, rheumatoid arthritis.

Table 2 HRs for incident fracture

	Number of events	Incidence*	Unadjusted HR	Unadjusted CI	Age/sex adjusted HR	Age/sex adjusted CI	Fully adjusted† HR	Fully adjusted† CI
All fractures								
Controls	49 168	92.18	REF		REF		REF	
PsA	575	99.23	1.09	1.00 to 1.18	1.14	1.05 to 1.24	1.16	1.06 to 1.27
Mild psoriasis	8470	92.38	1.01	0.98 to 1.03	1.09	1.07 to 1.12	1.07	1.05 to 1.10
Severe psoriasis	537	119.91	1.33	1.22 to 1.45	1.42	1.30 to 1.55	1.26	1.15 to 1.39
RA	3460	148.44	1.63	1.57 to 1.68	1.32	1.28 to 1.37	1.23	1.18 to 1.28
Hip fracture								
Controls	5930	10.71	REF		REF		REF	
PsA	54	8.97	0.86	0.66 to 1.12	1.27	0.97 to 1.66	1.17	0.86 to 1.59
Mild psoriasis	930	9.78	0.92	0.86 to 0.99	1.16	1.08 to 1.24	1.13	1.04 to 1.22
Severe psoriasis	55	11.77	1.17	0.90 to 1.53	1.69	1.29 to 2.20	1.21	0.88 to 1.66
RA	730	29.81	2.85	2.64 to 3.08	1.77	1.64 to 1.91	1.55	1.40 to 1.72
Vertebral fracture								
Controls	2009	3.62	REF		REF		REF	
PsA	20	3.32	0.94	0.60 to 1.46	1.06	0.69 to 1.65	1.07	0.66 to 1.72
Mild psoriasis	371	3.89	1.09	0.97 to 1.21	1.24	1.11 to 1.39	1.17	1.03 to 1.33
Severe psoriasis	32	6.85	2.02	1.42 to 2.87	2.35	1.66 to 3.33	2.23	1.54 to 3.22
RA	209	8.48	2.40	2.08 to 2.76	1.70	1.48 to 1.96	1.53	1.30 to 1.80

The fully adjusted models for each outcome were slightly different after employing a purposeful selection process. The variables contained within each model are specified as below.
*Incidence per 10 000 person-years.
†The all fracture model was adjusted for age, sex, cancer, atrial fibrillation, CKD, diabetes, COPD, liver disease, stroke, dementia, SSRI use, TCA use, antiepileptic use, PPI use, oral steroids, hormone treatment, ciclosporine, smoking and categorical BMI.
‡The hip fracture model was adjusted for age, sex, cancer, atrial fibrillation, CKD, CVD, diabetes, COPD, stroke, dementia, SSRI use, TCA use, antiepileptic use, oral steroids, hormone treatment, ciclosporine, smoking and categorical BMI.
§The vertebral fracture model was adjusted for age, sex, atrial fibrillation, diabetes, COPD, stroke, SSRI use, TCA use, PPI use, oral steroids, smoking and categorical BMI.
BMI, body mass index; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PPI, proton pump inhibitor; PsA, psoriatic arthritis; RA, rheumatoid arthritis; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.

adjusting for OP risk factors, patients with PsA and psoriasis had elevated risk for incident fracture: PsA 1.16 (95% CI 1.06 to 1.27), mild psoriasis 1.07 (1.05 to 1.10) and severe psoriasis HR 1.26 (1.15 to 1.39). Patients with mild psoriasis had an elevated risk for hip fracture 1.13 (1.04 to 1.22) and vertebral fracture (HR 1.17, 1.03 to 1.33). Patients with severe psoriasis had a substantially elevated risk for vertebral fracture: HR 2.23 (1.54 to 3.22). These results were robust to several sensitivity analyses (see online supplementary tables S3 and S4). Defining fracture by a code for fracture followed by a bisphosphonate prescription resulted in slight increases in the HR. The results did not substantially change when patients with a history of fracture were included.

DISCUSSION

In this study, we found patients with PsA and psoriasis had an increased prevalence of risk factors for OP and fracture (eg, diabetes, alcohol abuse, smoking, depression, antidepressant use, corticosteroids, methotrexate and ciclosporin).[24–29] Additionally, patients with PsA and psoriasis had an increased incidence of fracture compared with the general population by 7%–26%. The incidence of vertebral fracture was also increased in patients with severe psoriasis and while hip fracture was elevated in both psoriasis groups, it was only statistically significant in patients with mild psoriasis relative to matched controls after adjusting for risk factors for OP. We found that the risk for any fracture in patients with PsA and severe psoriasis was similar to RA. To our knowledge, these are the first population-based estimates of the risk for incident fracture and OP in patients with psoriasis and/or PsA and the first longitudinal cohort study to address this issue.

Strengths of this study include a large cohort of patients with an average of 6 years of follow-up; the use of THIN in which the exposures definitions (codes for PsA, RA, psoriasis) have been validated and fractures have been previously examined[13–17]; and the ability to adjust for other measured risk factors for OP, including concomitant medications, body mass index and smoking. Additionally, inclusion of a cohort of patients with RA for internal comparison provides validity to the results as our estimates for RA were similar to previous studies.[30–32] Similarly, the incidence of hip fracture among controls in our study was similar to population statistics in the UK (10.7 vs 10.3 per 10 000 person-years), further supporting the validity of our results.[33] Finally, the HRs were robust to numerous sensitivity analyses.

Our study also has limitations. There is a risk for misclassification of the outcome when using diagnosis codes to define an event rather than imaging. We addressed this through sensitivity analyses in which we changed the outcome definition; this did not significantly change the results. We also used a secondary definition for fracture in which we required a therapy for OP to address osteoporotic fractures. Vertebral fracture may be underdiagnosed and thus under recorded.[24] We conducted a sensitivity analysis to examine whether observation bias affected these results and found no difference when we only included patients in the study followed at least once yearly. Next, disease manifestations, disease activity and use of biological DMARDs are not available in THIN, and therefore we were unable to directly examine their effects on risk of OP and fracture. We were also unable to account for some lifestyle factors such as degree of immobility or laboratory parameters such as vitamin D. Finally, the relatively small number of patients with PsA and/or severe psoriasis may have resulted in insufficient power for some of the outcomes, resulting in wide CIs

that include 1.0 despite elevated point estimates (eg, for hip fracture among patients with severe psoriasis).[34]

In conclusion, fractures, in particular osteoporotic fractures, are a major health problem that results in poor outcomes and OP is largely underdiagnosed. We found that similar to PsA and psoriasis (both mild and severe) were associated with an increased risk for fractures. Screening and management of OP should still be considered for patients with psoriasis and PsA using guidelines available for the general population.[5 32]

Author affiliations
[1]Division of Rheumatology, Center for Clinical Epidemiology and Biostatistics, Center for Pharmacoepidemiology Research and Training, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA
[2]Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[3]Department of Dermatology, Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA
[4]Division of Rheumatology, Philadelphia Veterans Administration Medical Center, Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA
[5]Department of Dermatology, Center for Clinical Epidemiology and Biostatistics, Center for Pharmacoepidemiology Research and Training, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA
[6]Division of Rheumatology and Allergy/Immunology, Massachusetts General Hospital, Boston, Massachusetts, USA
[7]Division of Rheumatology, Landspitali University Hospital, University of Iceland, Reykjavik, Iceland

Acknowledgements The authors thank Yihui Jiang for administrative support.

Contributors AO and LH designed the study, performed the statistical analysis, wrote the first draft of the paper and integrated feedback. DS assisted in assembling the analytic dataset. All of the coauthors assisted in interpretation of the data and provided feedback on the manuscript draft. All authors have approved the final version of the manuscript.

Funding AO is supported by NIH by K23 AR063764. JT is supported by NIAMS K23AR068433. JFB is supported by a Veterans Affairs Clinical Science Research & Development Career Development Award (IK2 CX000955). JMG is supported by NIH/NIAMS K24 AR064310. Funds to extract the data relevant for this cohort study were provided by the Rheumatology Research Foundation Preceptorship Award.

Competing interests AO has received support from an investigator-initiated research grant from Pfizer and has consulted for Novartis and Pfizer and has received payment for continuing medical education work related to psoriatic arthritis. JT has an investigator-initiated research grant from Pfizer and has received payment for continuing medical education work related to psoriasis. JMG served as a consultant for Abbvie, Astrazeneca, Celgene, Coherus, Eli Lilly, Janssen Biologics (formerly Centocor), Sanofi, Merck, Novartis, Endo, Valeant and Pfizer, receiving honoraria; and received research grants (to the Trustees of the University of Pennsylvania) from Abbvie, Amgen, Eli Lilly, Janssen, Novartis, Regeneron and Pfizer; and received payment for continuing medical education work related to psoriasis.

Ethics approval This study used de-identified patient information, and was therefore eligible for institutional review board exemption. The Cegedim Scientific Review Committee also reviewed and approved the study protocol.

Provenance and peer review Not commissioned; externally peer reviewed.

Transparency AO affirms that the manuscript is an honest, accurate and transparent account of the study being reported; no aspects of the study have been omitted. There were no deviations from the original study plan.

Data sharing statement While we are unable to share the datasets, code lists are available on request.

REFERENCES

1 Wright NC, Looker AC, Saag KG, *et al*. The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine. *J Bone Miner Res* 2014;29:2520–6.
2 Singer A, Exuzides A, Spangler L, *et al*. Burden of illness for osteoporotic fractures compared with other serious diseases among postmenopausalwomen in the United States. *Mayo Clin Proc* 2015;90:53–62.
3 Goldring S. Inflammatory signaling induced bone loss. *Bone* 2015;80:143–9.
4 Del Puente A, Esposito A, Parisi A, *et al*. Osteoporosis and psoriatic arthritis. *J Rheumatol Suppl* 2012;89:36–8.
5 Ogdie A, Schwartzman S, Husni ME. Recognizing and managing comorbidities in psoriatic arthritis. *Curr Opin Rheumatol* 2015;27:118–26.
6 Dreiher J, Weitzman D, Cohen AD. Psoriasis and osteoporosis: a sex-specific association? *J Invest Dermatol* 2009;129:1643–9.
7 Keller JJ, Kang JH, Lin HC. Association between osteoporosis and psoriasis: results from the Longitudinal Health Insurance Database inTaiwan. *Osteoporos Int* 2013;24:1835–41.
8 Frediani B, Allegri A, Falsetti P, *et al*. Bone mineral density in patietns with psoriatic arthritis. *J Rheumatol* 2001;28:138–43.
9 Nolla JM, Fiter J, Rozadilla A, *et al*. Bone mineral density in patients with peripheral psoriatic arthritis. *Rev Rheum Engl Ed* 1999;66:457–61.
10 Riesco M, Manzano F, Font P, *et al*. Osteoporosis in psoriatic arthritis: an assessment of densitometry and fragility fractures. *Clin Rheumatol* 2013;32:1799–804.
11 Busquets N, Vaquero C, Moreno J, *et al*. Bone mineral density status and frequency of osteoporosis and clinical fractures in 155 patients with psoriaticarthritis followed in a university hospital. *Reumatol Clin* 2014;10:89–93.
12 Reddy S, Anandarajah A, Fisher M, *et al*. Comparative analysis of disease activity measures, use of biologic agents, body mass index, radiographic features, and bone density in psoriatic arthritis and rheumatoid arthritis patients followed in a large U. S. disease registry. *J Rheumatol Suppl* 2010;37:2566–72.
13 Seminara NM, Abuabara K, Shin DB, *et al*. Validity of The Health Improvement Network (THIN) for the study of psoriasis. *Br J Dermatol* 2011;164:602–9.
14 Ogdie A, Alehashemi S, Love TJ, *et al*. Validity of psoriatic arthritis and capture of disease modifying antirheumatic drugs in The Health Improvement Network. *Pharmacoepidemiol Drug Saf* 2014;23:918–22.
15 Rodríguez LA, Tolosa LB, Ruigómez A, *et al*. Rheumatoid arthritis in UK primary care: incidence and prior morbidity. *Scand J Rheumatol* 2009;38:173–7.
16 Watson DJ, Rhodes T, Cai B, *et al*. Lower Risk of Thromboembolic Cardiovascular Events with Naproxen Among Patients with Rheumatoid Arthritis. *Arch Intern Med* 2002;162:1105–10.
17 Watson DJ, Rhodes T, Guess HA. All-cause mortality and vascular events among patients with rheumatoid arthritis, osteoarthritis, or no arthritis in the UK General Practice Research Database. *J Rheumatol* 2003;30:1196–202.
18 Vosse D, Landewé R, van der Heijde D, *et al*. Ankylosing spondylitis and the risk of fracture: results from a large primary care-based nested case-control study. *Ann Rheum Dis* 2009;68:1839–42.
19 Gutiérrez L, Roskell N, Castellsague J, *et al*. Study of the incremental cost and clinical burden of hip fractures in postmenopausal women in the United Kingdom. *J Med Econ* 2011;14:99–107.
20 Gutiérrez L, Roskell N, Castellsague J, *et al*. Clinical burden and incremental cost of fractures in postmenopausal women in the United Kingdom. *Bone* 2012;51:324–31.
21 Misra D, Zhang Y, Peloquin C, *et al*. Incident long-term warfarin use and risk of osteoporotic fractures: propensity-score matched cohort of elders with new onset atrial fibrillation. *Osteoporos Int* 2014;25:1677–84.
22 Bursac Z, Gauss CH, Williams DK, *et al*. Purposeful selection of variables in logistic regression. *Source Code Biol Med* 2008;3:17.
23 Von Elm E, Altman DG, Egger M, *et al*. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Epidemiology* 2007;18:800–4.
24 Epidemiology of Osteoporosis. http://www.iofbonehealth.org/epidemiology (accessed 30 Jun 2015).
25 Wu Q, Qu W, Crowell MD, *et al*. Tricyclic antidepressant use and risk of fractures: a meta-analysis of cohort and case-control studies. *J Bone Miner Res* 2013;28:753–63.
26 Golob AL, Laya MB. Osteoporosis: screening, prevention and management. *Med Clin North Am* 2015;99:587–606.
27 Bernabei R, Martone AM, Ortolani E, *et al*. Screening, diagnosis and treatment of osteoporosis: a brief review. *Clin Cases Miner Bone Metab* 2014;11:201–7.
28 Collins GS, Mallett S, Altman DG. Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QFractureScores. *BMJ* 2011;342:d3651.
29 Vestergaard P, Rejnmark L, Mosekilde L. Methotrexate, azathioprine, cyclosporine and risk of fracture. *Calcif Tissue Int* 2006;79:69–75.
30 van Staa TP, Geusens P, Bijlsma JWJ, *et al*. Clinical assessment of the long-term risk of fracture in patients with rheumatoid arthritis. *Arthritis Rheum* 2006;54:3104–12.
31 Kim SY, Schneeweiss S, Liu J, *et al*. Risk of osteoporotic fracture in a large population-based cohort of patients with rheumatoid arthritis. *Arthritis Res Ther* 2010;12:R154.
32 Mullen MB, Saag KG. Evaluating and mitigating fracture risk in established rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2015;29:614–27.
33 National Clinical Guideline Centre (UK). *The management of hip fracture in adults*. London, UK: Royal College of Physicians, 2011.
34 Ogdie A, Troxel AB, Mehta NN, *et al*. Psoriasis and cardiovascular risk: strength in numbers part 3. *J Invest Dermatol* 2015;135:2148–50.
35 The Health Improvement Network: Our Data. http://csdmruk.cegedim.com/our-data/our-data.shtml (accessed 10 Jan 2017).

CONCISE REPORT

Performance of the ASAS classification criteria for axial and peripheral spondyloarthritis: a systematic literature review and meta-analysis

Alexandre Sepriano,[1,2] Roxana Rubio,[1] Sofia Ramiro,[1] Robert Landewé,[3,4] Désirée van der Heijde[1]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210747).

[1]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[2]NOVA Medical School, Universidade Nova de Lisboa, Lisboa, Portugal
[3]Department of Clinical Immunology & Rheumatology, Amsterdam Rheumatology & Clinical Immunology Center (ARC), Amsterdam, The Netherlands
[4]Zuyderland Medical Center, Heerlen, The Netherlands

Correspondence to
Dr Alexandre Rocha Sepriano, Leiden University Medical Center, Department of Rheumatology, Albinusdreef 2, P.O. Box 9600, Leiden 2300 RC, The Netherlands; alexsepriano@gmail.com

Received 31 October 2016
Revised 19 January 2017
Accepted 20 January 2017
Published Online First 8 February 2017

ABSTRACT

Objective To summarise the evidence on the performance of the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axial spondyloarthritis (axSpA) (also imaging and clinical arm separately), peripheral (p)SpA and the entire set, when tested against the rheumatologist's diagnosis ('reference standard').

Methods A systematic literature review was performed to identify eligible studies. Raw data on SpA diagnosis and classification were extracted or, if necessary, obtained from the authors of the selected publications. A meta-analysis was performed to obtain pooled estimates for sensitivity, specificity, positive and negative likelihood ratios, by fitting random effects models.

Results Nine papers fulfilled the inclusion criteria (N=5739 patients). The entire set of the ASAS SpA criteria yielded a high pooled sensitivity (73%) and specificity (88%). Similarly, good results were found for the axSpA criteria (sensitivity: 82%; specificity: 88%). Splitting the axSpA criteria in 'imaging arm only' and 'clinical arm only' resulted in much lower sensitivity (30% and 23% respectively), but very high specificity was retained (97% and 94% respectively). The pSpA criteria were less often tested than the axSpA criteria and showed a similarly high pooled specificity (87%) but lower sensitivity (63%).

Conclusions Accumulated evidence from studies with more than 5500 patients confirms the good performance of the various ASAS SpA criteria as tested against the rheumatologist's diagnosis.

INTRODUCTION

The Assessment of SpondyloArthritis international Society (ASAS) has developed and validated criteria (ASAS cohort) for spondyloarthritis (SpA), as well as for their subsets, axial (axSpA) and peripheral SpA (pSpA).[1 2] As in other rheumatic diseases,[3] in the absence of a 'true' gold-standard expert opinion has been used as an external 'anchor' to develop and test the SpA classification criteria. In the original validation studies, the ASAS criteria outperformed other classification criteria.

After their publication, the performance of the ASAS SpA criteria has been tested all over the world in different cohorts using the same approach. Some of these cohorts are expectedly similar to the ASAS cohort, while others differ (eg, setting, inclusion criteria, disease duration). Appropriate data pooling and exploring relevant between-study differences yield unique insights into the criteria performance and applicability in a broad population of patients.

The aim of this systematic literature review is to summarise the published data pertaining to the performance of the ASAS classification criteria for axSpA (also 'imaging arm' and 'clinical arm' separately), pSpA and the entire SpA set when tested against the rheumatologist's diagnosis.

METHODS

Literature search

The scope of the literature search was defined according to the PICO format (patients, intervention, comparator, outcomes; online supplementary table S1).[4] MEDLINE and EMBASE databases were searched without language restriction. Eligible studies were observational cohorts assessing the performance of the ASAS SpA criteria against the rheumatologist's diagnosis, published from March 2009 (date of the axSpA ASAS criteria release) up to August 2016. Studies in which the primary aim was not assessing the performance of the ASAS criteria but still provided enough data to allow such an analysis were also included. In order to retrieve additional references, abstracts from the American College of Rheumatology and European League Against Rheumatism annual conferences (2014 and 2015) were searched. Only studies with full text available were included, since abstracts neither provide appropriate detail for risk of bias (RoB) assessment nor appropriate data for analysis. Details on the search strategy are provided in online supplementary text 1.

Study selection, data extraction and assessment of risk of bias

Two reviewers (AS and RR) independently screened all titles and abstracts to identify eligible studies fulfilling the inclusion criteria followed by full-text review if appropriate (articles excluded and reason thereof in online supplementary table S2). Both reviewers independently extracted data on the studies' main characteristics, patient characteristics and disease characteristics, and criteria performance (ie, sensitivity, specificity, likelihood ratios of the ASAS criteria against the rheumatologist's diagnosis). Authors of the selected publications were contacted to obtain raw data (2×2 tables necessary for meta-analysis) on criteria performance, when this information was not available in the

To cite: Sepriano A, Rubio R, Ramiro S, *et al*. *Ann Rheum Dis* 2017;**76**:886–890.

publication. The same two reviewers independently assessed the RoB of each study using the Quality Assessment of Diagnostic Accuracy Studies 2 tool.[5] Disagreements were resolved by consensus, and a third review author was involved when necessary (DvdH).

Data analysis

Pooled sensitivity and specificity were estimated by random effects bivariate generalised linear mixed models. Parameter estimates from each model were used to derive the positive likelihood ratio (LR+) and negative LR (LR−) and 95% CIs. In case of limited data, two univariate random effects models were used by assuming no correlation between sensitivity and specificity.[6] Separate models were fit for the axSpA criteria, the pSpA criteria and the SpA criteria. The 'imaging arm' and the 'clinical arm' of the axSpA criteria were analysed separately using two approaches: (i) considering all patients that fulfil each arm irrespective of fulfilment of the other and (ii) considering patients that fulfil one arm exclusively.

A series of sensitivity analyses was performed (whenever possible and appropriate) to assess the effect of the following on the criteria performance: (i) target population (original validation study inclusion criteria vs different inclusion criteria); (ii) risk of bias (low vs high RoB); (iii) study's main aim (criteria performance assessment vs other); (iv) setting (hospital vs community) and (v) symptom duration (<2 years vs ≥2 years).

All analyses were performed in Stata V.12.1. The Cochrane Collaboration's Review Manager Software V.5.3 was used to build forest plots.

RESULTS

Of 1486 screened articles (after deduplication), 9 fulfilled the inclusion criteria (table 1).[1 2 7–13] All but one study were considered to be at low RoB (see online supplementary table S3). In total, 5739 patients (range: 157–1210) had been included, and 2936 (51.2%; range: 25.2%–69.4%) had been diagnosed by the rheumatologist as SpA.

Study populations

This literature review included the original studies in which the axSpA criteria and the pSpA criteria (also the entire set) were validated.[1 2] In addition, five studies assessed the ASAS axSpA criteria,[8–10 12 13] one study assessed the pSpA criteria[7] and one study the SpA criteria (providing separate data also for the axSpA and pSpA criteria).[11] Raw data on the criteria performance were obtained from all, except two studies.[12 13]

In table 1, main patient characteristics and disease characteristics per study are shown. The majority of the studies assessing the axSpA criteria had similar inclusion criteria compared with the original validation study.[8–10 12 13] However, in one study, inflammatory back pain was required, or otherwise patients had to have one additional SpA feature.[11]

Table 1 Main study characteristics

			Population (inclusion criteria)									
Study reference	Cohort	Sample size	Symptoms	Age symptoms onset (years)	Symptoms duration (years)	SpA* prevalence N (%)	Males (%)	Disease duration	HLA-B27 (%)	mNY (%)	MRI-SI (%)	Risk of bias
Rudwaleit *et al*[1]	ASAS	649	Any CBP (>3 months)	<45	No limit	391 (60.2)	52.4	6.1 (7.6) years	65.9	29.7	64.7†	Low
Rudwaleit *et al*[2]	ASAS	266	Arthritis/ enthesitis/ dactylitis	<45	No limit	176 (66.2)	63.1	10.3 (18.6) months	47.2	19.5	44.0†	Low
van den Berg *et al*[7]	EAC	302‡	Peripheral arthritis	NR	<2	76 (25.2)	48.7	22.8 (37.3) weeks	47.5	34.6	NR	Low
Moltó *et al*[8]	DECLIC	1210	Any CBP (>3 months)	<45	No limit	425 (35.1)	56.0	1.08 years (0.16, 3.90)§	60.1	49.2	25.2†	Low
van den Berg *et al*[9]	SPACE	157	Any CBP (>3 months)	<45	<2	65 (41.4)	48.3	13.4 (7.7) months	79.7	18.3	41.7¶	Low
Strand *et al*[10]	USA	816	Any CBP (>3 months)	<45	No limit	491 (60.2)	68.0	NR	NR	NR	NR	Low
Tomero *et al*[11]	ESPERANZA	775	IBP/ asymmetrical arthritis**	<45	<2	538 (69.4)	61.0	12.1 (6.8) months	56.0	19.0	24.0¶	Low
Lin *et al*[12]	China	867	Any CBP (>3 months)	<45	No limit	455 (52.5)	68.1	2.6 (3.2) years	72.3	NA	70.5¶	High
Deodhar *et al*[13]	PROSpA	697	Any CBP†† (>3 months)	<45	No limit	319 (45.8)	49.8	14.0 years	48.9	31.7	37.9¶	Low

For longitudinal studies, the baseline characteristics are shown. Characteristics are referring to patients with SpA according to the rheumatologist, except for the studies by van den Berg *et al*[7] (according to ASAS axSpA criteria) and Strand et al[10] (SpA and no-SpA).
*According to the rheumatologist's diagnosis (in the study by van den Berg et al,[7] prevalence of pSpA was calculated considering the 302 patients included in the analysis (prevalence in entire cohort: 76/2011=3.8%).
†Typical signs of active inflammation (no formal definition).
§Median (IQR).
‡Number of patients used in the analysis from a total of 2011 patients included in the cohort.
¶ASAS/Outcome Measures in Rheumatology (OMERACT) definition.
**In the absence of IBP or arthralgia only (without arthritis), one additional SpA feature required: psoriasis, inflammatory bowel disease, uveitis, radiographic sacroiliitis, positivity for HLA-B27 or a family history of SpA.
††And, ≥1 of the following: HLA-B27 positivity, current IBP and prior imaging (MRI or radiographic) evidence of sacroiliitis.
ASAS, Assessment of SpondyloArthritis international Society; axSpA, axial spondyloarthritis; CBP, chronic back pain; EAC, Early Arthritis Clinic; IBP, inflammatory back pain; mNY, modified New York criteria; NA, not applicable; NR, not reported; PROSpA, prevalence of axial SpA; pSpA, peripheral spondyloarthritis; SI, sacroiliitis; SpA, spondyloarthritis; SPACE, SpondyloArthritis Caught Early.

Two studies assessing the pSpA criteria used different inclusion criteria as compared with the ASAS cohort. In one study, only patients with peripheral arthritis were included (excluding those with only enthesitis or dactylitis),[7] while in another study patients had to have typical SpA arthritis (asymmetrical, and predominantly in lower limbs) or arthralgia associated with one additional SpA feature (not including enthesitis and dactylitis).[11]

Performance of the ASAS SpA classification criteria

The sensitivity and specificity of the various criteria for each individual study are shown in figure 1, and the results of the meta-analysis in table 2. The ASAS SpA criteria were assessed in two studies (N=1750) yielding a high pooled sensitivity and specificity (73%; 88%).[2 11]

Three studies (N=749) assessed the ASAS pSpA criteria.[2 7 11] Although specificity was consistently high (82%–90%; pooled: 87%), sensitivity was much lower in the two studies, with inclusion criteria differing from the original validation study (49%–56% vs 78%; pooled: 62%).

Seven studies, with 4990 patients in total, together generated a very high pooled sensitivity and specificity (82% and 87% respectively) for the axSpA criteria, with little variation across

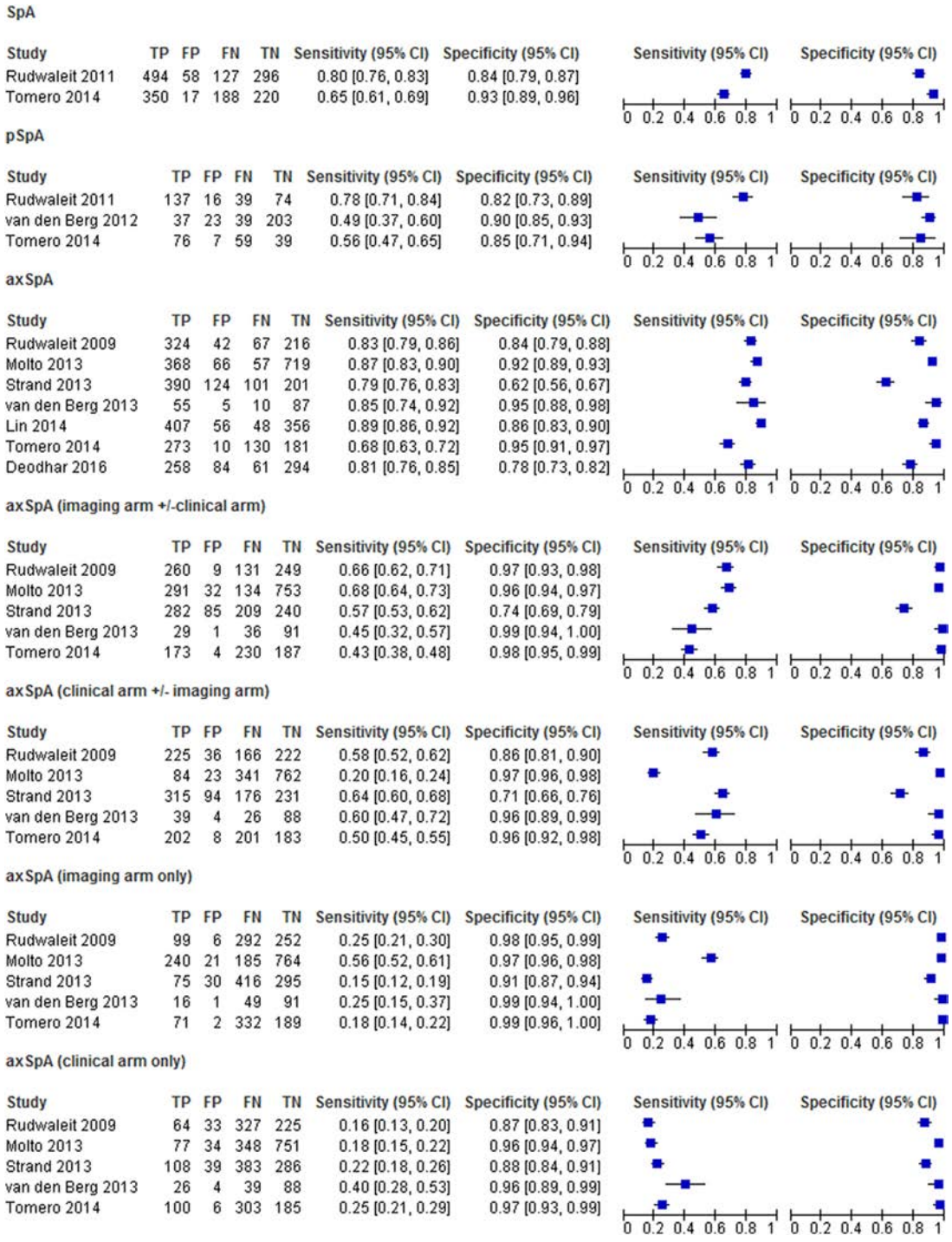


Figure 1 Performance of the ASAS SpA classification criteria across studies. ASAS, Assessment of SpondyloArthritis international Society; axSpA, axial spondyloarthritis; pSpA, peripheral spondyloarthritis; TP, true positives; FP, false positives; FN, false negatives; TN, true negatives.

Table 2 Results of the meta-analysis (pooled estimates)

	N patients (studies)	LR+ (95% CI)	LR− (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
ASAS SpA criteria*	1750 (2 studies[2 11])	6.3 (3.2 to 12.4)	0.31 (0.13 to 0.70)	0.73 (0.47 to 0.89)	0.88 (0.81 to 0.93)
ASAS pSpA criteria†	749 (3 studies[2 7 11])	4.7 (3.5 to 6.3)	0.43 (0.30 to 0.62)	0.62 (0.47 to 0.76)	0.87 (0.81 to 0.91)
ASAS axSpA criteria†	4990 (7 studies[1 8–13])	6.2 (3.7 to 10.5)	0.20 (0.16 to 0.27)	0.82 (0.77 to 0.86)	0.87 (0.78 to 0.92)
axSpA criteria† (imaging arm±clinical arm)	3426 (5 studies[1 8–11])	13.6 (4.8 to 38.7)	0.45 (0.37 to 0.56)	0.57 (0.47 to 0.66)	0.96 (0.88 to 0.99)
axSpA criteria† (clinical arm±imaging arm)	3426 (5 studies[1 8–11])	6.0 (2.9 to 12.4)	0.56 (0.43 to 0.72)	0.49 (0.34 to 0.64)	0.92 (0.82 to 0.96)
axSpA criteria† (imaging arm only)	3426 (5 studies[1 8–11])	9.6 (4.4 to 20.7)	0.76 (0.64 to 0.90)	0.26 (0.16 to 0.40)	0.97 (0.94 to 0.99)
axSpA criteria† (clinical arm only)	3426 (5 studies[1 8–11])	3.6 (2.0 to 6.4)	0.83 (0.75 to 0.91)	0.23 (0.17 to 0.29)	0.94 (0.90 to 0.96)

*Univariate random effects logistic regression.
†Bivariate random effects generalised mixed model.
ASAS, Assessment of SpondyloArthritis international Society; axSpA, axial spondyloarthritis; LR, likelihood ratio; pSpA, peripheral spondyloarthritis; SpA, spondyloarthritis.

studies.[1 8–13] The pooled sensitivity of the 'imaging arm'±'clinical arm' and 'clinical arm'±'imaging arm' was 57% and 49%, respectively (26% and 23% when considering patients fulfilling each arm exclusively). High estimates of pooled specificity were found for both 'arms', irrespective of the definition (range: 92%–97%). However, the LR+ of the 'imaging arm' only was higher as compared with the 'clinical arm' only (9.6 vs 3.6).

Sensitivity analyses

The ASAS axSpA criteria performed similarly well irrespective of the population in which they were applied, the setting, symptom duration, RoB and study's main aim (sensitivity (range): 78%–85%, specificity (range): 80%–93%; online supplementary table S4). Due to a scarcity of data, sensitivity analyses for the 'imaging arm' and 'clinical arm' of the axSpA criteria, the pSpA criteria and the SpA criteria could not be performed.

DISCUSSION

Pooled data from eight cohorts (including more than 5500 patients) confirm the good performance of the various ASAS SpA classification criteria as tested against the rheumatologist's diagnosis. This review confirms that splitting the 'arms' of the axSpA criteria results in loosing sensitivity while retaining specificity, which indicates that the full set of axSpA criteria is the preferred set.

While the pooled specificity for both the axSpA criteria and pSpA criteria was similarly high (87% for both), the pooled sensitivity for the pSpA criteria was much lower than that for the axSpA criteria (62% vs 82%). This difference may be explained by restrictive inclusion criteria. Unlike the ASAS cohort, the Early Arthritis Clinic cohort only included patients with arthritis, and not those with dactylitis only or enthesitis only.[7] Similar 'restrictions' were seen in the ESPERANZA cohort.[11] The low sensitivity found in these studies suggests that both enthesitis and dactylitis are considered by the rheumatologists as fitting the pattern of pSpA, which adds to the credibility of the ASAS pSpA criteria (that include these presentations).

Sensitivity analyses have shown the 'robustness' of the axSpA criteria when applied in different settings (hospital and community), in patients with short (<2 years) and long (≥2 years) symptom duration and in different populations.

Not surprisingly, the splitting of the axSpA criteria into two 'arms' compromised sensitivity, but retained (very high) specificity, if patients that fulfil each 'arm' irrespective of fulfilment of the other were considered, and if those that fulfil one 'arm' exclusively were analysed. The larger LR+ for the 'imaging arm' as compared with the 'clinical arm' reflects the rheumatologist's reliance on positive imaging findings. The prospective validation of the ASAS criteria against the rheumatologist's diagnosis after >4 years of follow-up in the ASAS cohort has shown that both 'arms' still properly discriminate between axSpA and no-axSpA.[14] Another prospective study has also suggested the arms' low specificity when tested against radiographic sacroiliitis (modified New York criteria) after 8 years of follow-up ('imaging arm': 22%; 'clinical arm': 56%), but the setting in this study was a prognostic rather than a diagnostic setting, and figures are difficult to interpret.[15]

In conclusion, the ASAS axSpA and pSpA criteria have shown to perform well in patients included in several cohorts all over the world, as assessed by rheumatologists. This review does not give resolution to the applicability of the ASAS classification criteria in primary care, since such a setting had not been tested. It is important to realise that the criteria's performance depends entirely on the prevalence of SpA in the underlying population (pretest likelihood).

Acknowledgements The authors thank the authors of the included papers for providing raw data. They also thank Yemisi Takwoingi, co-convenor for the Cochrane Screening and Diagnostic Test Methods, for the support in statistical analysis.

Contributors Study concept and design: AS, SR, RL and DvdH. Data collection: AS and RR. Statistical analysis and data interpretation: AS, SR, RR, RL and DvdH. All authors revised the manuscript critically for important intellectual content and gave final approval of the version to be published. AS prepared the first version of the manuscript.

Funding AS received a research grant from *Fundação para a Ciência e Tecnologia* (grant number: SFRH/BD/108246/2015).

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Rudwaleit M, van der Heijde D, Landewé R, *et al*. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–83.
2 Rudwaleit M, van der Heijde D, Landewé R, *et al*. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.
3 Aggarwal R, Ringold S, Khanna D, *et al*. Distinctions between diagnostic and classification criteria? *Arthritis Care Res (Hoboken)* 2015;67:891–7.
4 Sackett D, WS R, Rosenberg W, *et al*. *Evidence-based medicine: how to practice and teach EBM*. London: Churchill Livingstone, 1997.
5 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
6 Takwoingi Y, Guo B, Riley RD, *et al*. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Stat Methods Med Res* Epub ahead of print: 26 Jun 2015. doi:10.1177/0962280215592269
7 van den Berg R, van Gaalen F, van der Helm-van Mil A, *et al*. Performance of classification criteria for peripheral spondyloarthritis and psoriatic arthritis in the Leiden Early Arthritis cohort. *Ann Rheum Dis* 2012;71:1366–9.

8 Moltó A, Paternotte S, Comet D, *et al*. Performances of the Assessment of SpondyloArthritis International Society axial spondyloarthritis criteria for diagnostic and classification purposes in patients visiting a rheumatologist because of chronic back pain: results from a multicenter, cross-sectional study. *Arthritis Care Res (Hoboken)* 2013;65:1472–81.
9 van den Berg R, de Hooge M, van Gaalen F, *et al*. Percentage of patients with spondyloarthritis in patients referred because of chronic back pain and performance of classification criteria: experience from the Spondyloarthritis Caught Early (SPACE) cohort. *Rheumatology (Oxford)* 2013;52:1492–9.
10 Strand V, Rao SA, Shillington AC, *et al*. Prevalence of axial spondyloarthritis in United States rheumatology practices: Assessment of SpondyloArthritis International Society criteria versus rheumatology expert clinical diagnosis. *Arthritis Care Res (Hoboken)* 2013;65:1299–306.
11 Tomero E, Mulero J, de Miguel E, *et al*. Performance of the Assessment of Spondyloarthritis International Society criteria for the classification of spondyloarthritis in early spondyloarthritis clinics participating in the ESPERANZA programme. *Rheumatology (Oxford)* 2014;53:353–60.
12 Lin Z, Liao Z, Huang J, *et al*. Evaluation of Assessment of Spondyloarthritis International Society classification criteria for axial spondyloarthritis in Chinese patients with chronic back pain: results of a 2-year follow-up study. *Int J Rheum Dis.* 2014;17:782–9.
13 Deodhar A, Mease PJ, Reveille JD, *et al*. Frequency of Axial Spondyloarthritis Diagnosis Among Patients Seen by US Rheumatologists for Evaluation of Chronic Back Pain. *Int J Rheum Dis* 2016;68:1669–76.
14 Sepriano A, Landewé R, van der Heijde D, *et al*. Predictive validity of the ASAS classification criteria for axial and peripheral spondyloarthritis after follow-up in the ASAS cohort: a final analysis. *Ann Rheum Dis* 2016;75:1034–42.
15 Aydin SZ, Maksymowych WP, Bennett AN, *et al*. Validation of the ASAS criteria and definition of a positive MRI of the sacroiliac joint in an inception cohort of axial spondyloarthritis followed up for 8 years. *Ann Rheum Dis* 2012;71:56–60.

EXTENDED REPORT

Hydroxychloroquine inhibits proinflammatory signalling pathways by targeting endosomal NADPH oxidase

Nadine Müller-Calleja,[1,2] Davit Manukyan,[1,2] Antje Canisius,[1] Dennis Strand,[3] Karl J Lackner[1]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210012).

[1]Institute of Clinical Chemistry and Laboratory Medicine, University Medical Center Mainz, Mainz, Germany
[2]Center for Thrombosis and Hemostasis, University Medical Center Mainz, Mainz, Germany
[3]Department of Medicine 1, University Medical Center Mainz, Mainz, Germany

Correspondence to
Dr Karl J Lackner, Institute of Clinical Chemistry and Laboratory Medicine, University Medical Center Mainz, Langenbeckstrasse 1, D-55131 Mainz, Germany; karl.lackner@unimedizin-mainz.de

Received 3 June 2016
Revised 28 September 2016
Accepted 5 November 2016
Published Online First 30 November 2016

ABSTRACT

Objectives Hydroxychloroquine (HCQ) has been used for decades to treat patients with rheumatic diseases, for example, systemic lupus erythematosus (SLE), rheumatoid arthritis or the antiphospholipid syndrome (APS). We hypothesise that HCQ might target endosomal NADPH oxidase (NOX), which is involved in the signal transduction of cytokines as well as antiphospholipid antibodies (aPL).

Methods For in vitro experiments, monocytic cells were stimulated with tumour necrosis factor α (TNFα), interleukin-1β (IL-1β) or a human monoclonal aPL and the activity of NOX was determined by flow cytometry. The expression of genes known to be induced by these stimuli was quantified by quantitative reverse transcription PCR. Live cell imaging was performed by confocal laser scanning microscopy. Finally, the effects of HCQ on NOX-induced signal transduction were analysed in an in vivo model of venous thrombosis.

Results HCQ strongly reduces or completely prevents the induction of endosomal NOX by TNFα, IL-1β and aPL in human monocytes and MonoMac1 cells. As a consequence, induction of downstream genes by these stimuli is reduced or abrogated. This effect of HCQ is not mediated by direct interference with the agonists but by inhibiting the translocation of the catalytic subunit of NOX2 (gp91phox) into the endosome. In vivo, HCQ protects mice from aPL-induced and NOX2-mediated thrombus formation.

Conclusions We describe here a novel mechanism of action of HCQ, that is, interference with the assembly of endosomal NOX2. Since endosomal NOX2 is involved in many inflammatory and prothrombotic signalling pathways, this activity of HCQ might explain many of its beneficial effects in rheumatic diseases including the APS.

INTRODUCTION

Antimalarial drugs and in particular chloroquine and hydroxychloroquine (HCQ) have been used for decades in the treatment of rheumatic and autoimmune diseases.[1–4] The efficacy of HCQ in patients with systemic lupus erythematosus (SLE) or mild rheumatoid arthritis has been well documented even though the effects are moderate compared with more potent immunosuppressant drugs.[5–7] HCQ reduces the risk for thromboembolic events in patients with SLE and positive for antiphospholipid antibodies (aPL).[8–10] Accordingly, it has been recommended recently that patients with SLE and aPL should be treated with HCQ as a prophylactic measure.[11]

While several potential mechanisms of action have been identified, the exact mode of action of HCQ is still under debate.[3 5] Furthermore, it is not known, if the efficacy of HCQ in different rheumatic and autoimmune diseases is mediated by different mechanisms. The previously described actions of HCQ include reduction of cytokine production,[12–14] inhibition of immune effector cells,[3] inhibition of platelet function,[15] protection of the cell surface from external disturbances,[16] competitive binding to nucleic acid ligands of toll-like receptors (TLRs),[17] interference with lysosomal function[3 5] and reduction of leakage of lysosomal enzymes.[3] While some of these quite heterogeneous effects associated with HCQ might be related to each other, no unifying mechanism of action of HCQ has been found.

Due to high affinity of HCQ to the lysosomal/endosomal compartment, a potentially relevant target for HCQ might be endosomal NADPH oxidase (NOX). This enzyme complex is involved in numerous proinflammatory signalling cascades.[18] In particular, signalling of tumour necrosis factor α (TNFα) via TNF-receptor 1 (TNFR1) and interleukin-1β (IL-1β) via IL-1R are mediated in part by uptake of the ligand-receptor complexes into the endosome, activation of endosomal NOX and generation of superoxide and subsequently other reactive oxygen species (ROS).[19–21] Inhibition of endosomal NOX massively reduces downstream activation of NFκB via these pathways. It should be noted though that signalling still proceeds with reduced intensity indicating that the endosomal route accounts for only part of the cytokine effects.[18]

We have recently shown that aPL also induce endosomal NOX in monocytes and endothelial cells providing a novel mechanism of action of aPL.[22–24] Indeed, ability of certain aPL to induce venous thrombosis in vivo depends on NOX2.[25] In our hands, aPL induced a more rapid and intense activation of endosomal NOX than TNFα or IL-1β reflected in a much stronger ROS signal. Thus, aPL are another activator of endosomal NOX.

Since inhibition of TNFα and also type 1 interleukins and their signalling pathways has been shown to be therapeutically effective in rheumatoid arthritis[2] and probably other autoimmune diseases including SLE[26 27] and HCQ shows therapeutic

To cite: Müller-Calleja N, Manukyan D, Canisius A, *et al*. *Ann Rheum Dis* 2017;**76**:891–897.

benefits in these diseases as well as in antiphospholipid syndrome (APS), we hypothesised that endosomal NOX might be targeted by HCQ. If this hypothesis is correct, it might help to explain many of the observed in vivo effects of HCQ and the efficacy of HCQ in several autoimmune diseases.

MATERIALS AND METHODS

Human aPL

The human monoclonal aPL HL5B has been previously described in detail.[28–30] It was cloned from a patient with primary APS and has numerous somatic mutations. It binds to phospholipids in a cofactor-independent manner. An unspecific IgG was generated by the same method. IgG fractions of patients and controls were obtained as described previously (see online supplementary table S1).[24 31] All antibody and IgG preparations had <0.1 U endotoxin/mL as determined by limulus amoebocyte assay. All patients provided informed consent according to the ethical guidelines following the Declaration of Helsinki. Collection and use of blood samples has been approved by the ethics committee of the State Medical Association of Rheinland-Pfalz.

Cell culture and stimulation of cells

MonoMac1 (MM1) cells were maintained in RPMI-1640 medium supplemented with 10% fetal calf serum, L-glutamine and sodium pyruvate. Monoclonal aPL (100 ng/mL), IgG fractions (100 µg/mL), TNFα or IL-1β (both ebioscience, 10 ng/mL) were added as indicated to 0.5×10^6 cells/mL. HCQ (Sigma Aldrich, usually 10 µM or as indicated) was added 15 min before the stimuli. To analyse gene expression, RNA was isolated and quantitative reverse transcriptase PCR was performed as previously described.[31] Similarly, isolation and culture of human monocytes and mouse monocytes from C57BL/6J mice and gp91phox$^{-/-}$ mice on the same genetic background have been described.[23 24]

Flow cytometric detection of cellular ROS formation

The detection of endosomal ROS generation was performed by use of the fluorogenic reagent OxyBURST Green H_2HFF-BSA.[32] Cells were kept in phosphate-buffered saline containing 1 mM Ca^{2+}, 1.5 mM Mg^{2+} and 5.5 mM glucose for 2 hours before being incubated in 10 µg/mL H_2HFF-BSA for 2 min with or without HCQ (10 µM) at 37°C. Thereafter, cells were stimulated with agonists as indicated. ROS-induced cellular fluorescence was analysed by flow cytometry.

Confocal laser scanning microscopy

Microscopy was performed with a Zeiss LSM 710 NLO confocal laser scanning microscope and a 1.4 Oil Dic M27 63×plan apochromat objective (Zeiss). To show the intracellular localisation of NOX2 and TLR8 on stimulation with agonists, cells were fixed and stained with fluorescence-labelled anti-TLR8 or anti-gp91phox. Labelled antibodies against calnexin and EEA-1 were used as markers for the endoplasmic reticulum or early endosomes as described by Latz.[33] For live cell microscopy, cells were cultured in RPMI without phenol red, stimulated as indicated and imaged directly in chambers maintained at 37°C (Nunc). Monoclonal aPL (HL5B) and control IgG were labelled with fluorescein isothiocyanate (FITC) by the use of a standard FITC Antibody Labeling Kit (Thermo Fisher Scientific).

In vivo thrombosis model and intravital microscopy

The in vivo model of thrombus formation in the inferior vena cava (IVC) used in this study has been previously described in detail.[25 34] It is based on severe flow reduction in the IVC. Briefly, in anesthetised mice, a median laparotomy was performed and the IVC was exposed. A permanent narrowing ligature was applied exactly below the left renal vein. Human monoclonal aPL HL5B (1 µg) was injected via a jugular catheter 1 hour before flow reduction in the IVC. HCQ (10 µg) was injected intravenously 2 hours before HL5B injection as indicated. Acridine orange was injected (20 µg) intravenously to stain circulating leucocytes in vivo.[34] Murine platelets were isolated from whole blood as described[34] and labelled with 20 µg/mL rhodamine B. Thrombus formation was observed and quantified by high-speed real-time intravital fluorescence video microscopy (BX51WI; Olympus).

Statistics

All numerical data are shown as mean±SD. Normal distribution was confirmed using Shapiro-Wilk test. Statistical analyses of the data were performed by Student's t-test for normally distributed data and Wilcoxon test for not normally distributed data. p Values <0.05 were considered statistically significant for single testing.

RESULTS

HCQ blocks endosomal NOX activation

Several stimuli including TNFα, IL-1β and certain aPL, for example, the human monoclonal HL5B, induce endosomal NOX2 in monocytic cells followed by downstream effects.[22 23 32] HCQ has high affinity to acidic compartments, that is, lysosomes and endosomes. Therefore, we analysed its influence on endosomal NOX activation in MM1 cells by flow cytometry. As expected, all three stimuli induced significant ROS production with HL5B giving rise to the most rapid ROS production (figure 1A–C). Niflumic acid (NFA), an inhibitor of chloride channel 3 (ClC3), completely blocked cellular ROS production. Since NFA has been shown previously to selectively prevent superoxide generation by endosomal NOX,[32] this confirms that ROS induced by TNFα, IL-1β and HL5B is generated by endosomal NOX. HCQ inhibited ROS production in a dose-dependent manner (figure 1D–F). All effects observed in MM1 cells could be reproduced in primary human monocytes (see online supplementary figure S1).

Gene induction in MM1 cells

To analyse whether blockade of endosomal ROS production by HCQ also affects known cellular responses to TNFα, IL-1β and aPL, we determined the effect of HCQ on the induction of genes known to be rapidly and strongly induced by these three agonists. TNFα induces its own secretion via NFκB.[35 36] IL-1β stimulates IL-8 release in a NFκB-dependent manner.[37] We have previously shown that HL5B rapidly induces tissue factor (TF) expression.[23] Again the effects of HCQ were compared with those of NFA. Both substances were added 15 min before the respective agonists. As shown in figure 2, NFA and HCQ were equally efficacious in suppressing gene induction by the three agonists. While the effects of IL-1β and aPL on IL-8 and TF were completely blocked at 10 µM HCQ, it appeared that blockade of TNFα was not fully complete. At 3 µM HCQ, inhibition of the agonists was slightly less but still significant. These effects of HCQ could be confirmed on the protein level (data not shown). In addition HCQ blocks aPL-induced translocation of TLR8 to the endosome (see online supplementary

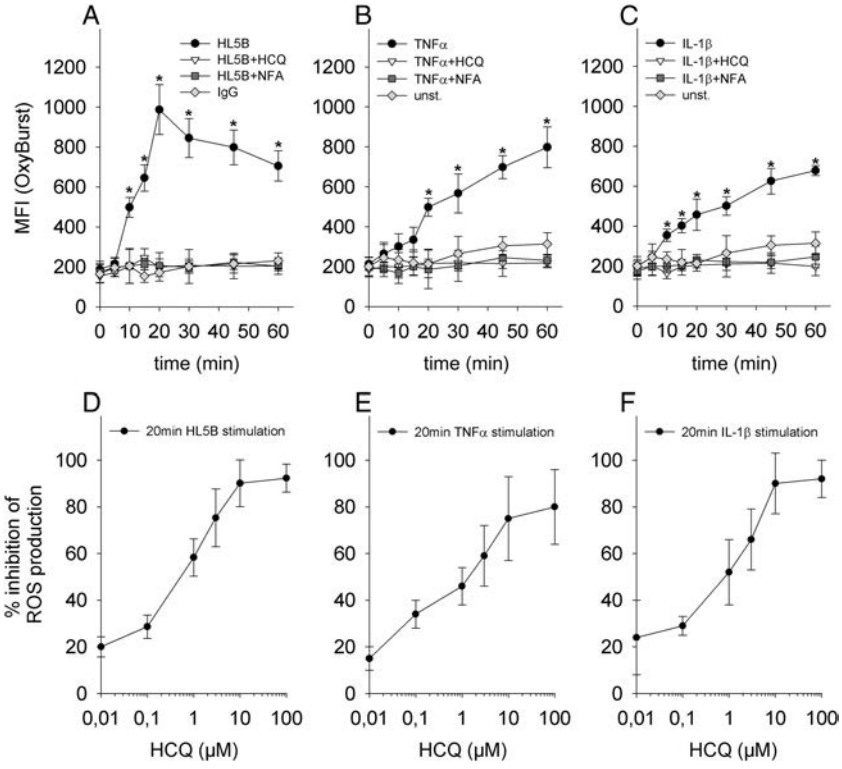


Figure 1 Effect of hydroxychloroquine (HCQ) on superoxide generation. (A–C) MonoMac1 (MM1) cells were loaded with the reactive oxygen species (ROS) sensitive dye OxyBurst before stimulation. Cells were stimulated for up to 60 min with 100 ng/mL HL5B or IgG (A), 10 ng/mL tumour necrosis factor α (TNFα) (B) or 10 ng/mL interleukin-1β (IL-1β) (C) either alone or together with HCQ (10 μM) or niflumic acid (0.1 mM) and absolute fluorescence recorded at the indicated time points. Data are from six independent experiments measured in duplicate. *p<0.01 agonist versus agonist+HCQ. (D–F) Dose–response curves of HCQ effects on ROS production induced by HL5B (D), TNFα (E) or IL-1β (F) in MM1 cells. An almost maximal response is achieved at 10 μM HCQ.

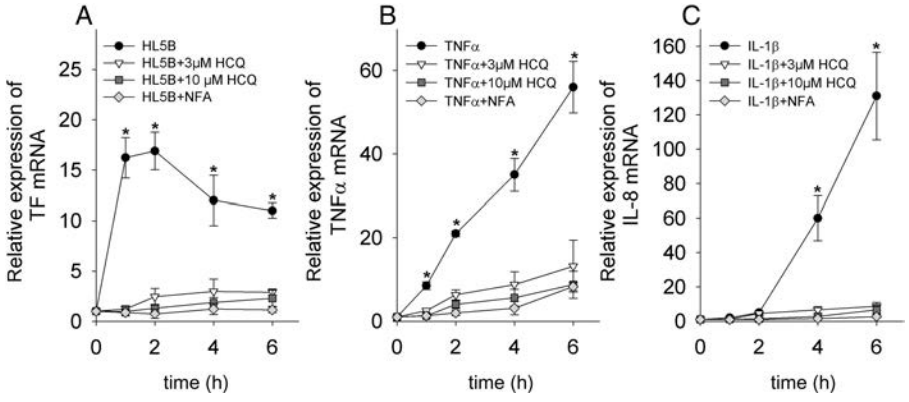


Figure 2 Effect of hydroxychloroquine (HCQ) on gene induction. MonoMac1 cells were stimulated for up to 6 hours as described in figure 1. Relative expression of tissue factor mRNA (A), tumour necrosis factor α (TNFα) mRNA (B) and interleukin-8 (IL-8) mRNA (C) was normalised to IgG stimulated (A) or unstimulated cells (B+C) and β-actin expression. Data are from six independent experiments (three experiments for 3 μM HCQ and niflumic acid) measured in duplicate. *p<0.05 agonist versus agonist+HCQ.

figure S2). Our data also confirm that under the cell culture conditions used all effects of the three agonists strongly depend on endosomal NOX.

HCQ does not affect TNFα and IL-1β signalling in NOX2-deficient cells

To assess the contribution of NOX2 inhibition on the overall effect of HCQ, we compared monocytes from gp91phox$^{-/-}$ and wild type C57BL/6J mice. Responses to TNFα and IL-1β were detectable but reduced by 70%–90% in gp91phox$^{-/-}$ mice. Addition of 10 μM HCQ did not have any effect on the residual response in these cells, while it reduced the response in wild type cells to the level of NOX2-deficient cells (figure 3). This suggests that the effect of HCQ on TNFα and IL-1β signalling is more or less exclusively mediated via its effect on endosomal NOX2.

Mechanism of NOX inhibition

Activation of endosomal NOX by the three agonists depends on the transport of ligand-receptor complexes or aPL into the endosome. HCQ is a lysosomotropic agent that can possibly prevent clathrin-dependent endocytosis.[38 39] We, therefore, analysed if HCQ has any influence on aPL internalisation. The effect of HCQ on aPL endocytosis was analysed by confocal microscopy using FITC-labelled HL5B. As shown before,[22] HL5B is rapidly internalised into the endosomal route as shown by overlap with LysoTracker, a marker for endosomes and lysosomes. There were no discernible effects of HCQ on the pattern of intracellular distribution of HL5B and LysoTracker, providing evidence that HCQ has no effect on internalisation and endosomal accumulation of the monoclonal aPL HL5B (see online supplementary figure S3).

HCQ blocks gp91phox translocation to early endosomes

NOX2 is a membrane-bound enzyme complex of six subunits that converts molecular oxygen to superoxide. In resting cells, the catalytic units gp91phox and p22phox, collectively referred to as flavocytochrome b558, are integrated in membranes, while the other components remain soluble. On stimulation, the cytosolic proteins migrate to the membrane to assemble the active

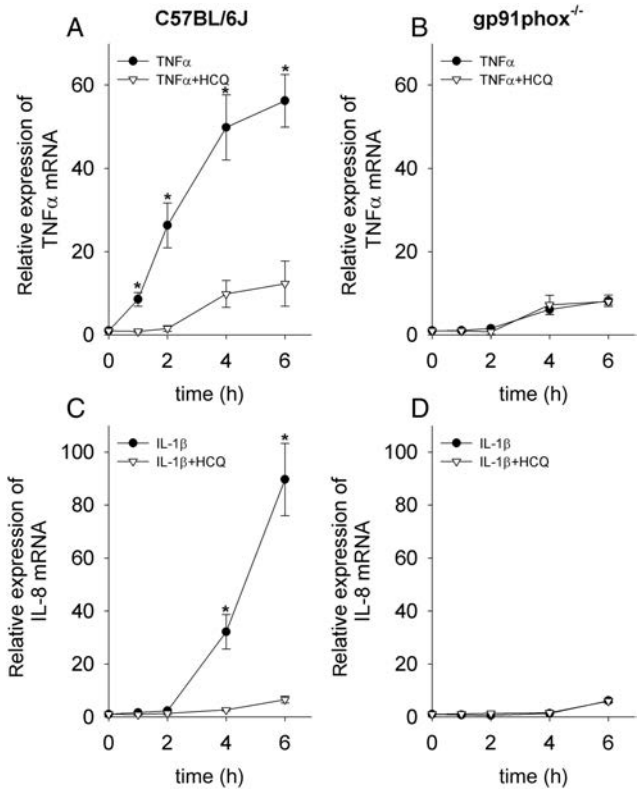


Figure 3 Hydroxychloroquine (HCQ) has no effect on residual tumour necrosis factor α (TNFα) and interleukin-1β (IL-1β) activity in gp91phox-deficient mice. CD115+ monocytes isolated from C57BL/6 wild type mice (A and C) or gp91phox$^{-/-}$ mice (B and D) were stimulated for up to 6 hours with 10 ng/mL of TNFα (A and B) or IL-1β (C and D). Relative expression of TNFα mRNA (A and B) and IL-8 mRNA (C and D) was normalised to unstimulated cells and β-actin expression. Induction of indicated genes was dramatically reduced in NOX2 deficient animals. However, residual cytokine induction was not affected by addition of HCQ. Data are from three independent experiments measured in duplicate. *p<0.01 agonist versus agonist +HCQ.

oxidase.[40–42] In monocytes and macrophages, the mature flavocytochrome is found primarily in the plasma membrane but also in the endocytic recycling compartment.[43]

To visualise the localisation of NOX2, MM1 cells were incubated with a fluorescently labelled antibody against gp91phox and a marker for early endosomes (EEA1). In resting cells or cells incubated with control IgG, there was no detectable overlap between gp91phox and EEA1 (figure 4A–C, left panels). Incubation with HL5B, TNFα or IL-1β rapidly induced movement of gp91phox to the endosome as shown by almost complete colocalisation of gp91phox with EEA1. This is the evidence that the three agonists induce assembly of the active multiprotein complex of NOX2 in endosomes. Addition of HCQ to the cell culture completely blocked the movement of gp91phox to the endosome (figure 4). In fact, the distribution of gp91phox after stimulation with HL5B, TNFα or IL-1β in the presence of HCQ was similar to unstimulated cells. In contrast to HCQ, NFA did not prevent movement of gp91phox to the endosome. As previously described, NFA prevents ROS production by blockade of ClC3. Taken together, these data reveal that TNFα, IL-1β and HL5B induce the translocation of gp91phox to early endosomes in monocytic cells, where the active NOX2 enzyme complex is assembled. This early step in endosomal NOX2 activation is blocked by HCQ.

HCQ prevents signalling induced by IgG fractions from patients with APS

We have previously shown that IgG isolated from 15 patients with APS without exception induced the same signalling cascade as HL5B.[24] To exclude that the effect of HCQ was limited to our monoclonal aPL, we confirmed our data using these 15 IgG fractions of patients with APS (see online supplementary table S1). HCQ prevented the increase of endosomal ROS production induced by APS-IgG (figure 5A). As expected, HCQ also almost completely prevented the induction of TF mRNA expression induced by APS-IgG (figure 5B). TF mRNA expression after incubation of MM1 cells with IgG of patients with APS in the presence of HCQ was only slightly higher than mRNA levels of MM1 cells incubated with IgG of 15 healthy control donors (male/female: 5/10; range: 21–72 years).

HCQ prevents HL5B-induced thrombosis in a mouse model

To show that the in vitro effects of HCQ are also relevant in vivo, we chose an in vivo thrombosis model, previously described. This mouse model is based on flow reduction in the IVC and detection of thrombus formation by intravital microscopy. In this mouse model, HL5B and a similar monoclonal aPL, RR7F, massively accelerate venous thrombus formation. This effect is fully dependent on activation of NOX2 as it is absent in gp91phox-deficient mice.[25] Pretreatment of the mice with HCQ significantly reduced HL5B-induced venous thrombus formation (figure 6A, B). Together with the in vitro data, this is a strong evidence that inhibition of activation of endosomal NOX by HCQ is also relevant in vivo.

DISCUSSION

We present a novel mechanism of action how HCQ exerts its therapeutically relevant anti-inflammatory and antithrombotic effects in vitro and in vivo. HCQ blocks a signalling pathway common to TNFα, IL-1β and aPL, which depends on activation of endosomal NOX2 and leads to proinflammatory and procoagulant cellular responses. HCQ concentrations effective in vitro (1–10 μM) are comparable with therapeutic plasma levels of HCQ (3±2 μM). The single dose of HCQ applied in the in vivo mouse model (approx. 0.4 mg/kg body weight) was even lower than therapeutic doses in humans. Interestingly, the major effect of HCQ is the prevention of agonist-induced gp91phox translocation into the endosome. Accordingly, no induction of endosomal NOX2 activity (ie, ROS production) is detectable. This specific extraendosomal effect of HCQ makes a non-specific action by increasing endosomal/lysosomal pH unlikely. As a lysosomotropic weak base, HCQ is rapidly protonated, thereby increasing the pH of endolysosomal vesicles. This may inhibit lysosomal enzymes that require an acidic pH, and prevent fusion of endosomes and lysosomes.[44–46] However, a pH optimum for NOX2 between 6.5 and 8.0 has been reported.[47] Moreover, it has been shown that NOX2 itself contributes to limiting the acidification of early endosomes in dendritic cells.[48] Thus, endosomal NOX2 activity does not depend on acidic pH. Furthermore, it has been shown that 4 μM HCQ affects lysosomal pH minimally[17] and in our hands, the labelling pattern of the pH sensitive LysoTracker at 10 μM HCQ is indistinguishable from control.

NOX2 is a membrane-bound enzyme complex generating superoxide. It is made up of the membrane-bound catalytic core (flavocytochrome b558) consisting of the integral proteins p22phox and gp91phox and four cytosolic subunits: p47phox, p67phox, p40phox and Rac1/2.[49–51] On stimulation, the

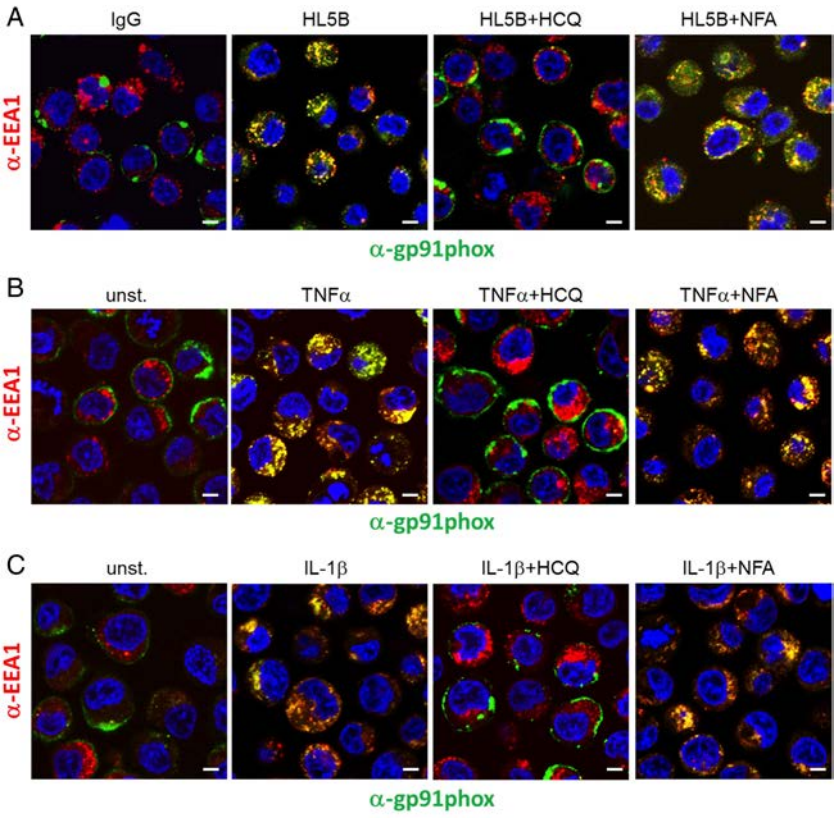


Figure 4 Hydroxychloroquine (HCQ) blocks gp91phox translocation into the endosome. MonoMac1 cells were stimulated for 30 min with (A) HL5B or control IgG (100 ng/mL), (B) tumour necrosis factor α (TNFα) (10 ng/mL) or (C) interleukin-1β (IL-1β) (10 ng/mL). HCQ (10 μM) or niflumic acid (NFA) (0.1 mM) were added as indicated. After fixation, cells were stained with anti-(α)-gp91phox (green), anti-(α)-EEA1 (red) and 4,6-diamidino-2-phenylindole (DAPI) (blue) and visualised by confocal laser scanning microscopy. In unstimulated or IgG stimulated cells, gp91phox is detected at the cell membrane. Incubation with HL5B, TNFα or IL-1β leads to translocation of gp91phox to the endosome as shown by colocalisation with EEA1. This translocation was prevented by addition of HCQ but not by addition of NFA.

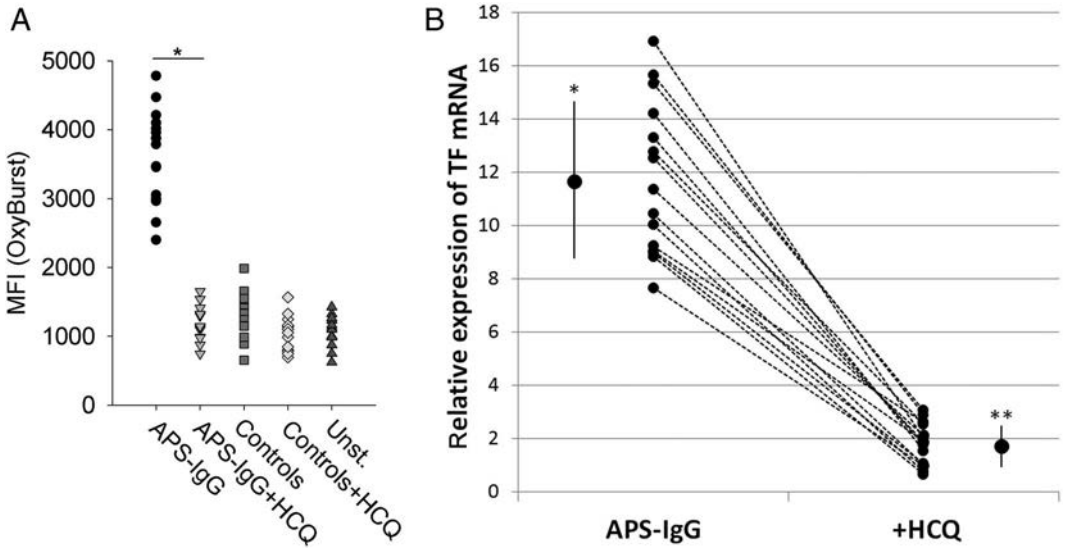


Figure 5 Hydroxychloroquine (HCQ) prevents reactive oxygen species (ROS) release and tissue factor (TF) induction induced by antiphospholipid syndrome (APS) IgG fractions. MonoMac1 cells were stimulated with IgG fractions (100 μg/mL) of 15 patients with APS and 15 healthy control donors in the absence or presence of HCQ (10 μM) as indicated. (A) Endosomal superoxide production after 20 min of stimulation was detected by flow cytometry using OxyBurst as ROS sensitive dye (*p<0.01). (B) Relative expression of TF (after 1 hour incubation) was normalised to cells stimulated with IgG fractions of 15 healthy control donors (relative mean expression is 1) and to the expression of β-actin. The large symbols indicate mean±SD of the respective normalised mRNA expression. *p<0.0001 versus control IgG and versus HCQ; **p<0.005 versus control IgG.

soluble factors and flavocytochrome b558 assemble the active enzyme complex and superoxide is released.[52 53] Thus, NOX2 activity is mainly controlled at the level of the correct and timely assembly of preformed subunits, which is prevented by HCQ.

This mechanism of NOX inhibition by HCQ is obviously different from NFA, an inhibitor of ClC3.[32] It has been proposed that ClC3 provides charge neutralisation of NOX-generated electron flux.[32] Alternatively, ClC3 may serve as an ion channel for superoxide itself and provide a means to leave the endosome. Thus, NFA does not interfere with NOX2 assembly but inhibits ROS production by the enzyme complex.

In any case, both HCQ and NFA prevent all cellular events downstream of NOX2. This includes induction of TF, TNFα

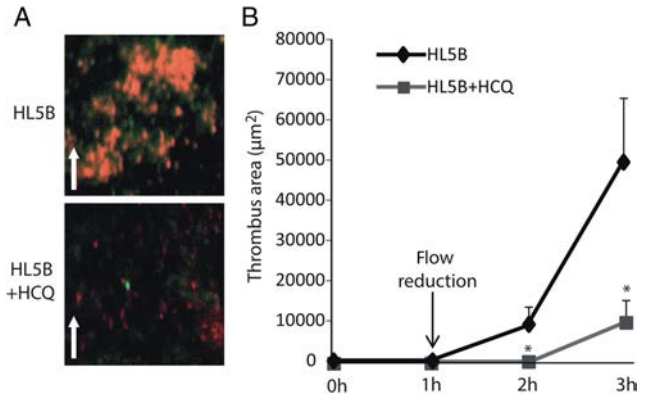


Figure 6 Hydroxychloroquine (HCQ) inhibits thrombus formation induced by HL5B. C57BL/6J mice were infused with HL5B 1 hour before flow reduction in the inferior vena cava. HCQ (10 µg) was infused 2 hours before application of HL5B as indicated. Thrombus formation was visualised by intravital microscopy. (A) Representative microscopic image of the vena cava vessel wall in animals treated with HL5B (upper panel) and HL5B+HCQ (lower panel). Direction of flow is indicated by an arrow. Platelets are labelled red/orange and leucocytes green. (B) Quantitative analysis of thrombus area in HL5B (n=7) and in HL5B+HCQ (n=6) infused mice (*p<0.05).

and IL-8 as well as translocation of TLR8 to the endosome. While the induction of cytokines depends on activation of NFκB by ROS, translocation of TLR8 to the endosome induced by aPL depends on superoxide generation but not on NFκB.[22] Thus, endosomal ROS-induced downstream effects are not limited to induction of NFκB and genes regulated by this nuclear factor.

To translate our in vitro data to the in vivo situation, we made use of a mouse model of venous thrombosis. We have shown that activation of endosomal NOX2 by aPL greatly accelerates thrombus formation in this model. While wild type C57BL/6J mice rapidly develop venous thrombi when exposed to aPL, gp91phox-deficient mice are protected.[25] We show now that HCQ can also prevent aPL-induced thrombus formation in this in vivo model. Thus, our data provide evidence that inhibition of NOX2 activation by HCQ is also a relevant pharmacologic action of this drug in the intact organism.

Inhibition of endosomal NOX2 can explain several well-established effects of HCQ, that is, reduction of cytokine production and plasma concentrations[12–14] or inhibition of different immune effector cells.[3] Reduced activity of endosomal NOX2 on autocrine or paracrine activation of immune cells by TNFα and IL-1β will lead to reduced production of their target cytokines. Since signalling by TNFα and IL-1β is not exclusively mediated by the endosomal pathway, inhibition by HCQ is most likely incomplete and the effects of HCQ, for example, on signalling by TNFα will be less pronounced compared with effects of direct TNF inhibitors such as adalimumab or etanercept. This is exactly what one might expect from the clinical efficacy profile of HCQ, which is most useful in milder cases of rheumatic diseases. On the other hand, HCQ is able to reduce the effects of other agonists besides TNFα. The relative importance of these pathways with respect to the therapeutic effects of HCQ needs further investigations. This applies particularly to IL-1β, which has been discussed as a potential therapeutic target in rheumatic diseases.

In the case of aPL, we have previously shown that signalling by aPL of similar specificity as our monoclonal aPL HL5B is mediated exclusively via endosomal NOX2. We have elucidated the cellular signalling events induced by HL5B and patient IgG fractions in detail previously.[22–25] Their effects are completely blocked by HCQ in vitro and in vivo. In particular, the prevention of thrombus induction in vivo by HCQ is of major relevance. We propose that this effect of HCQ provides an explanation for its beneficial role in the prevention of thromboembolic events in patients with aPL.[10 11] However, it should be noted that there are other aPL with specificity for β2-glycoprotein I (β2GPI), which have also been shown to be pathogenic in vitro and in vivo.[54] These aPL most likely induce cellular responses by other signal transduction pathways, which are probably dependent on formation of a complex of β2GPI/anti-β2GPI. Rand *et al* have shown that HCQ can disintegrate these complexes and prevent dislocation of annexin A5 from the cell surface. They proposed that this might be an explanation for the protective effect of HCQ.[16] Their data imply that the interaction of HCQ with cofactor-independent aPL analysed by us and anti-β2GPI may be quite different. Interestingly, there is some evidence that endosomal uptake of anti-β2GPI may be required for their pathogenic effects.[55] However, the relevance of endosomal NOX2 in this process has not yet been analysed.

In conclusion, we present a novel mechanism how HCQ exerts its well-established anti-inflammatory and thromboprophylactic effects. Since signalling endosomes serve as physical platforms for crosstalk between different signalling pathways,[56] this might explain the apparently heterogeneous therapeutic profile of HCQ.

Acknowledgements We thank Dr Steffen Lorenz from the Imaging Core Facility Mainz for providing assistance with confocal microscopy.

Contributors NM-C designed the study, performed experiments, wrote the manuscript; DM performed in vivo experiments; AC performed experiments; DS performed confocal microscopy; KJL designed the study, wrote the manuscript.

Funding This work was supported by the Federal Ministry of Education and Research (BMBF 01EO1003).

Competing interests None.

Ethics approval Ethikkommission der Landesärztekammer Rheinland-Pfalz.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Ruiz-Irastorza G, Ramos-Casals M, Brito-Zeron P, *et al*. Clinical efficacy and side effects of antimalarials in systemic lupus erythematosus: a systematic review. *Ann Rheum Dis* 2010;69:20–8.
2 Smolen JS, Landewé R, Breedveld FC, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73:492–509.
3 Rainsford KD, Parke AL, Clifford-Rashotte M, *et al*. Therapy and pharmacological properties of hydroxychloroquine and chloroquine in treatment of systemic lupus erythematosus, rheumatoid arthritis and related diseases. *Inflammopharmacology* 2015;23:231–69.
4 Tang C, Godfrey T, Stawell R, *et al*. Hydroxychloroquine in lupus: emerging evidence supporting multiple beneficial effects. *Intern Med J* 2012;42:968–78.
5 Costedoat-Chalumeau N, Dunogué B, Morel N, *et al*. Hydroxychloroquine: a multifaceted treatment in lupus. *Presse Med* 2014;43(Pt 2):e167–180.
6 A randomized study of the effect of withdrawing hydroxychloroquine sulfate in systemic lupus erythematosus. The Canadian Hydroxychloroquine Study Group. *N Engl J Med* 1991;324:150–4.
7 Katz SJ, Russell AS. Re-evaluation of antimalarials in treating rheumatic diseases: re-appreciation and insights into new mechanisms of action. *Curr Opin Rheumatol* 2011;23:278–81.
8 Kaiser R, Cleveland CM, Criswell LA. Risk and protective factors for thrombosis in systemic lupus erythematosus: results from a large, multi-ethnic cohort. *Ann Rheum Dis* 2009;68:238–41.
9 Fangtham M, Petri M. 2013 update: Hopkins lupus cohort. *Curr Rheumatol Rep* 2013;15:360.
10 Belizna C. Hydroxychloroquine as an anti-thrombotic in antiphospholipid syndrome. *Autoimmun Rev* 2015;14:358–62.

11 Erkan D, Aguiar CL, Andrade D, *et al*. 14th International Congress on Antiphospholipid Antibodies: task force report on antiphospholipid syndrome treatment trends. *Autoimmun Rev* 2014;13:685–96.
12 Sperber K, Quraishi H, Kalb TH, *et al*. Selective regulation of cytokine secretion by hydroxychloroquine: inhibition of interleukin 1 alpha (IL-1-alpha) and IL-6 in human monocytes and T cells. *J Rheumatol* 1993;20:803–8.
13 van den Borne BE, Dijkmans BA, de Rooij HH, *et al*. Chloroquine and hydroxychloroquine equally affect tumor necrosis factor-alpha, interleukin 6, and interferon-gamma production by peripheral blood mononuclear cells. *J Rheumatol* 1997;24:55–60.
14 Tishler M, Yaron I, Shirazi I, *et al*. Hydroxychloroquine treatment for primary Sjögren's syndrome: its effect on salivary and serum inflammatory markers. *Ann Rheum Dis* 1999;58:253–6.
15 Achuthan S, Ahluwalia J, Shafiq N, *et al*. Hydroxychloroquine's efficacy as an antiplatelet agent study in healthy volunteers: a proof of concept study. *J Cardiovasc Pharmacol Ther* 2015;20:174–80.
16 Rand JH, Wu XX, Quinn AS, *et al*. Hydroxychloroquine protects the annexin A5 anticoagulant shield from disruption by antiphospholipid antibodies: evidence for a novel effect for an old antimalarial drug. *Blood* 2010;115:2292–9.
17 Kuznik A, Bencina M, Svajger U, *et al*. Mechanism of endosomal TLR inhibition by antimalarial drugs and imidazoquinolines. *J Immunol* 2011;186:4794–804.
18 Spencer NY, Engelhardt JF. The basic biology of redoxosomes in cytokine-mediated signal transduction and implications for disease-specific therapies. *Biochemistry* 2014;53:1551–64.
19 Li Q, Spencer NY, Oakley FD, *et al*. Endosomal Nox2 facilitates redox-dependent induction of NF-kappaB by TNF-alpha. *Antioxid Redox Signal* 2009;11:1249–63.
20 Oakley FD, Smith RL, Engelhardt JF. Lipid rafts and caveolin-1 coordinate interleukin-1beta (IL-1beta)-dependent activation of NFkappaB by controlling endocytosis of Nox2 and IL-1beta receptor 1 from the plasma membrane. *J Biol Chem* 2009;284:33255–64.
21 Tchikov V, Bertsch U, Fritsch J, *et al*. Subcellular compartmentalization of TNF receptor-1 and CD95 signaling pathways. *Eur J Cell Biol* 2011;90:467–75.
22 Prinz N, Clemens N, Strand D, *et al*. Antiphospholipid antibodies induce translocation of TLR7 and TLR8 to the endosome in human monocytes and plasmacytoid dendritic cells. *Blood* 2011;118:2322–32.
23 Prinz N, Clemens N, Canisius A, *et al*. Endosomal NADPH-oxidase is critical for induction of the tissue factor gene in monocytes and endothelial cells. Lessons from the antiphospholipid syndrome. *Thromb Haemost* 2013;109:525–31.
24 Müller-Calleja N, Köhler A, Siebald B, *et al*. Cofactor-independent antiphospholipid antibodies activate the NLRP3-inflammasome via endosomal NADPH-oxidase: implications for the antiphospholipid syndrome. *Thromb Haemost* 2015;113:1071–83.
25 Manukyan D, Müller-Calleja N, Jäckel S, *et al*. Cofactor-independent human antiphospholipid antibodies induce venous thrombosis in mice. *J Thromb Haemost* 2016;14:1011–20.
26 Kirou KA, Gkrouzman E. Anti-interferon alpha treatment in SLE. *Clin Immunol* 2013;148:303–12.
27 Aringer M, Smolen JS. Therapeutic blockade of TNF in patients with SLE-promising or crazy? *Autoimmun Rev* 2012;11:321–5.
28 Lackner KJ, von Landenberg C, Barlage S, *et al*. Analysis of prothrombotic effects of two human monoclonal IgG antiphospholipid antibodies of apparently similar specificity. *Thromb Haemost* 2000;83:583–8.
29 Prinz N, Häuser F, Lorenz M, *et al*. Structural and functional characterization of a human IgG monoclonal antiphospholipid antibody. *Immunobiology* 2011;216:145–51.
30 von Landenberg C, Lackner KJ, von Landenberg P, *et al*. Isolation and characterization of two human monoclonal anti-phospholipid IgG from patients with autoimmune disease. *J Autoimmun* 1999;13:215–23.
31 Döring Y, Hurst J, Lorenz M, *et al*. Human antiphospholipid antibodies induce TNFalpha in monocytes via Toll-like receptor 8. *Immunobiology* 2010;215:230–41.
32 Miller FJ Jr, Filali M, Huss GJ, *et al*. Cytokine activation of nuclear factor kappa B in vascular smooth muscle cells requires signaling endosomes containing Nox1 and ClC-3. *Circ Res* 2007;101:663–71.
33 Latz E, Schoenemeyer A, Visintin A, *et al*. TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nat Immunol* 2004;5:190–8.
34 von Brühl ML, Stark K, Steinhart A, *et al*. Monocytes, neutrophils, and platelets cooperate to initiate and propagate venous thrombosis in mice in vivo. *J Exp Med* 2012;209:819–35.
35 Coward WR, Okayama Y, Sagara H, *et al*. NF-kappa B and TNF-alpha: a positive autocrine loop in human lung mast cells? *J Immunol* 2002;169:5287–93.
36 Guergnon J, Chaussepied M, Sopp P, *et al*. A tumour necrosis factor alpha autocrine loop contributes to proliferation and nuclear factor-kappaB activation of Theileria parva-transformed B cells. *Cell Microbiol* 2003;5:709–16.
37 Kim GY, Lee JW, Ryu HC, *et al*. Proinflammatory cytokine IL-1beta stimulates IL-8 synthesis in mast cells via a leukotriene B4 receptor 2-linked pathway, contributing to angiogenesis. *J Immunol* 2010;184:3946–54.
38 Salisbury JL, Condeelis JS, Satir P. Role of coated vesicles, microfilaments, and calmodulin in receptor-mediated endocytosis by cultured B lymphoblastoid cells. *J Cell Biol* 1980;87:132–41.
39 Wang LH, Rothberg KG, Anderson RG. Mis-assembly of clathrin lattices on endosomes reveals a regulatory switch for coated pit formation. *J Cell Biol* 1993;123:1107–17.
40 Nauseef WM. Assembly of the phagocyte NADPH oxidase. *Histochem Cell Biol* 2004;122:277–91.
41 IJdo JW, Mueller AC. Neutrophil NADPH oxidase is reduced at the anaplasma phagocytophilum phagosome. *Infect Immun* 2004;72:5392–401.
42 DeLeo FR, Burritt JB, Yu L, *et al*. Processing and maturation of flavocytochrome b558 include incorporation of heme as a prerequisite for heterodimer assembly. *J Biol Chem* 2000;275:13986–93.
43 Casbon AJ, Allen LA, Dunn KW, *et al*. Macrophage NADPH oxidase flavocytochrome B localizes to the plasma membrane and Rab11-positive recycling endosomes. *J Immunol* 2009;182:2325–39.
44 Dunmore BJ, Drake KM, Upton PD, *et al*. The lysosomal inhibitor, chloroquine, increases cell surface BMPR-II levels and restores BMP9 signalling in endothelial cells harbouring BMPR-II mutations. *Hum Mol Genet* 2013;22:3667–79.
45 Ohkuma S, Poole B. Fluorescence probe measurement of the intralysosomal pH in living cells and the perturbation of pH by various agents. *Proc Natl Acad Sci USA* 1978;75:3327–31.
46 Solomon VR, Lee H. Chloroquine and its analogs: a new promise of an old drug for effective and safe cancer therapies. *Eur J Pharmacol* 2009;625:220–33.
47 Morgan D, Cherny VV, Murphy R, *et al*. The pH dependence of NADPH oxidase in human eosinophils. *J Physiol* 2005;569(Pt 2):419–31.
48 Mantegazza AR, Savina A, Vermeulen M, *et al*. NADPH oxidase controls phagosomal pH and antigen cross-presentation in human dendritic cells. *Blood* 2008;112:4712–22.
49 Chanock SJ, el Benna J, Smith RM, *et al*. The respiratory burst oxidase. *J Biol Chem* 1994;269:24519–22.
50 Quinn MT, Gauss KA. Structure and regulation of the neutrophil respiratory burst oxidase: comparison with nonphagocyte oxidases. *J Leukoc Biol* 2004;76:760–81.
51 El-Benna J, Dang PM, Gougerot-Pocidalo MA, *et al*. Phagocyte NADPH oxidase: a multicomponent enzyme essential for host defenses. *Arch Immunol Ther Exp (Warsz)* 2005;53:199–206.
52 Groemping Y, Rittinger K. Activation and assembly of the NADPH oxidase: a structural perspective. *Biochem J* 2005;386(Pt 3):401–16.
53 Vignais PV. The superoxide-generating NADPH oxidase: structural aspects and activation mechanism. *Cell Mol Life Sci* 2002;59:1428–59.
54 Giannakopoulos B, Krilis SA. The pathogenesis of the antiphospholipid syndrome. *N Engl J Med* 2013;368:1033–44.
55 Brandt KJ, Fickentscher C, Boehlen F, *et al*. NF-κB is activated from endosomal compartments in antiphospholipid antibodies-treated human monocytes. *J Thromb Haemost* 2014;12:779–91.
56 Pálfy M, Reményi A, Korcsmáros T. Endosomal crosstalk: meeting points for signaling pathways. *Trends Cell Biol* 2012;22:447–56.

EXTENDED REPORT

An expanded population of pathogenic regulatory T cells in giant cell arteritis is abrogated by IL-6 blockade therapy

Chie Miyabe, Yoshishige Miyabe, Klemen Strle, Nancy D Kim, John H Stone, Andrew D Luster, Sebastian Unizony

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210070).

Division of Rheumatology, Allergy and Immunology, Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

Correspondence to
Dr Sebastian Unizony or Dr Andrew D Luster, Division of Rheumatology, Allergy and Immunology, Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, Harvard Medical School, 55 Fruit Street Boston, MA 02114, USA; sunizony@mgh.harvard.edu or aluster@mgh.harvard.edu

ADL and SU contributed equally.

Received 16 June 2016
Revised 5 September 2016
Accepted 9 November 2016
Published Online First 16 December 2016

ABSTRACT

Objectives Randomised-controlled trials have recently proven the efficacy of the interleukin (IL)-6 receptor antagonist tocilizumab (TCZ) in giant cell arteritis (GCA). However, the mechanism of action of IL-6 blockade in this disease is unknown. Moreover, the role of regulatory T (Treg) cells in the pathogenesis of GCA remains underexplored. Given the plasticity of Tregs and the importance of IL-6 in their biology, we hypothesised that TCZ might modulate the Treg response in GCA. We therefore characterised the Treg compartment of patients with GCA treated with TCZ.

Methods We classified 41 patients with GCA into three groups: active disease (aGCA, n=11), disease remission on corticosteroids (rGCA-CS, n=19) and disease remission on TCZ (rGCA-TCZ, n=11). Healthy controls (HCs) were included for comparison. We determined the frequency, phenotype and function of peripheral blood Tregs.

Results Patients with aGCA demonstrated a hypoproliferating Treg compartment enriched in IL-17-secreting Tregs (IL-17$^+$Tregs). Tregs in patients with aGCA disproportionally expressed a hypofunctional isoform of Foxp3 that lacks exon 2 (Foxp3Δ2). Foxp3Δ2-expressing Tregs coexpressed CD161, a marker commonly associated with the Th17 linage, significantly more often than full-length Foxp3-expressing Tregs. Compared with those of HCs, GCA-derived Tregs demonstrated impaired suppressor capacity. Treatment with TCZ, in contrast to CS therapy, corrected the Treg abnormalities observed in aGCA. In addition, TCZ treatment increased the numbers of activated Tregs (CD45RA$^-$Foxp3^{high}) and the Treg expression of markers of trafficking (CCR4) and terminal differentiation (CTLA-4).

Conclusions TCZ may exert its therapeutic effects in GCA by increasing the proliferation and activation of Tregs, and by reverting the pathogenic Treg phenotype seen during active disease.

INTRODUCTION

Giant cell arteritis (GCA) is the most frequent primary vasculitis in Western countries.[1] The main histopathological feature of the disease comprises a granulomatous inflammatory process rich in CD4$^+$T cells and macrophages that involves large-sized and medium-sized arteries.[1] Most patients develop relapsing courses despite prolonged treatments with corticosteroid (CS), which invariably lead to drug-related toxicity.[2] Agents that maintain disease remission and spare the use of CS are therefore the greatest unmet need for this patient population.[3–6]

An imbalance among CD4$^+$T helper (Th)1, Th17 and regulatory T (Treg) cells is thought to contribute to the pathogenesis of GCA.[7–9] Patients with new-onset disease demonstrate Th1 and Th17 cell infiltrates in their arteries and an expansion of these cell subsets in peripheral blood.[7–9] Conversely, decreased numbers of Tregs in the peripheral circulation are found in patients with GCA, regardless of the state of disease activity.[8 9] Although the Th17 axis is sensitive to CS treatment,[7–10] some reports suggest that the abnormalities described in both the Th1 and Treg subsets are resistant to CS therapy,[7 8] thereby accounting for the high relapse rate in GCA following CS tapering.

The interleukin (IL)-6 pathway is a novel target in GCA. Patients with GCA demonstrate increased IL-6 RNA expression within inflamed arteries[11 12] and elevated IL-6 protein levels in the peripheral blood during active disease.[13] Recently, two randomised controlled trials showed that tocilizumab (TCZ), a monoclonal antibody against the IL-6 receptor (IL-6R), is effective in maintaining disease remission in absence of CS.[14 15] However, the mechanism of action of IL-6 signalling blockade in GCA remains unknown.

Considerable phenotypical and functional plasticity exists within the Treg and the Th17 cell subsets.[16] Th17 cells and Tregs develop from a common naïve CD4$^+$T cell precursor under the influence of transforming growth factor-β (TGF-β).[17] In the presence of proinflammatory mediators (eg, IL-6 or IL-21), TGF-β-stimulated CD4$^+$T cells differentiate into Th17 cells, whereas in the absence of an inflammatory microenvironment these TGF-β-stimulated precursors are induced to become Tregs.[18] Furthermore, under specific circumstances, fully differentiated Tregs may lose their suppressive function and become IL-17-producing cells[16] (eg, 'pathogenic Tregs'[19 20] and exFoxp3 Th17 cells).[21]

One mechanism regulating the divergent fates between Tregs and Th17 cells involves the molecular antagonism of RAR-related orphan receptor (ROR) γt (RORC in humans) by Foxp3 through the domain encoded by the exon 2 of the FOXP3 gene.[22] Tregs that express a spliced variant of Foxp3 lacking exon 2 (Foxp3Δ2) are less suppressive,[23] and more likely to become IL-17 producing Tregs.

To cite: Miyabe C, Miyabe Y, Strle K, *et al*. *Ann Rheum Dis* 2017;**76**:898–905.

Increased numbers of Foxp3Δ2$^+$Tregs have been reported in anti-neutrophil cytoplasmic antibody (ANCA)-associated vasculitis.[24] It is not known, however, whether this abnormality is also present in patients with GCA. In addition, cells that express both Foxp3 and IL-17 have been detected in inflamed GCA arteries,[10] but whether this cell population is present in peripheral circulation, and most importantly, whether disease treatment reverts the pathogenic phenotype of those Tregs has been insufficiently studied.

We aimed to characterise the regulatory CD4$^+$T cell compartment in peripheral blood of patients with GCA and to investigate the effects of IL-6R blockade therapy with TCZ on the frequency, phenotype and function of those cells.

MATERIALS AND METHODS

Study population

We evaluated 41 patients with GCA in a cross-sectional study. Patients with GCA were classified into one of three categories based on disease activity and treatment: active disease (aGCA, n=11), disease remission on CS monotherapy (rGCA-CS, n=19) and disease remission on TCZ therapy (rGCA-TCZ, n=11). Among the subjects with aGCA, three had new-onset disease and eight were sampled during a disease relapse. We also evaluated samples from 10 healthy controls (HCs). Upon diagnosis, all patients had been treated with CS according to the standard of care for GCA.[1] Patients in the TCZ group (rGCA-TCZ group) received their IL-6R blockade therapy because of relapsing disease or prohibitive CS-related toxicity during previous treatment courses. Once on TCZ, patients underwent a prednisone taper of variable rate, but generally faster than the standard of care in order to ameliorate or prevent CS-related toxicity. Other clinical information is provided in the online supplementary text.

Cell isolation, culture and flow cytometry

CD4$^+$T cells were purified (>90% purity) from whole blood using RosetteSep CD4$^+$enrichment antibody cocktail (StemCell Technologies) according to manufacturer's instructions. Cells were labelled with Pacific Blue-conjugated anti-CD4, fluorescein isothiocyanate-conjugated anti-CD45RA, phycoerythrin (PE)-Cy7-conjugated anti-CCR4, PE-conjugated anti-CTLA4, allophycocyanin (APC)-Cy7-conjugated anti-IL-17A, PE-Cy7-conjugated anti-CD25, PE-conjugated anti-Ki67, APC-conjugated anti-CD161 (BioLegend); Alexa Fluor 488-conjugated anti-Foxp3 and Alexa Fluor 700-conjugated anti-Foxp3. Foxp3Δ2 was detected using clone PCH101 (eBiocience) that recognises the N-terminus portion of the protein and clone 150D (BioLegend) that recognises exon 2.[24 25] Data were acquired on a LSRFortessa cell analyser (BD biosciences) and analysed with FlowJo software.

Treg suppression assays

CD4$^+$CD25$^+$Tregs were isolated from a pool of CD4$^+$T cells using CD25 MicroBeads (Miltenyi Biotec). CD4$^+$CD25$^-$ conventional T cells were incubated for 10 min at 37°C in 10 µM carboxyfluorescein N-hydroxysuccinimidyl ester (CFSE) (Invitrogen), washed with phosphate buffered saline containing 2% fetal calf serum (FCS), and resuspended in complete Roswell Park Memorial Institute (RPMI) medium. CFSE-labelled CD4$^+$CD25$^-$cells (1×10^5) were co-incubated with varying concentrations of autologous CD4$^+$CD25$^+$Tregs to create conventional T cell to Treg ratios of 8:1, 4:1, 2:1 and 1:1. Cultures were stimulated for 4 days with Treg Suppression Inspector (Miltenyi Biotec), or left unstimulated. Proliferation of conventional CD4$^+$T cells was measured by assessing CFSE dilution by flow cytometry.

Statistical analysis

Categorical variables were compared between groups using Fisher's exact test. Continuous variables were compared between groups using paired and unpaired Student's t-test, Mann-Whitney test, analysis of variance or Kruskal-Wallis test as appropriate. In order to account for confounders on the number of specific CD4$^+$T cell subsets (eg, CS dose) we used linear regression. Statistical significance cut-off was 0.05. p Values were two-sided. Stata V.13 (StataCorp LP) was used for all analyses.

RESULTS

Baseline characteristics of patients with GCA and HCs

The baseline characteristics of patients with GCA and HCs are shown in table 1. There were no significant differences among patient groups (aGCA, rGCA-TCZ and rGCA-CS) with regard to demographic features, disease type or disease duration. The mean daily dose of prednisone at the time of blood sampling was 15.7 mg in patients in the rGCA-CS group, 0.2 mg in patients in the rGCA-TCZ group and 8.0 mg in patients in the aGCA group (p=0.02). Patients in the rGCA-TCZ group had received TCZ for a median period of 18 months. The HCs were younger than the patients with GCA (mean age 59 years vs 72 years; p<0.01), but there were no other important differences.

TCZ increases the frequency of activated Tregs

We first measured the population of Tregs defined as CD4$^+$T cells expressing Foxp3 and found no significant differences among groups (see online supplementary figure S1A, B). We then classified Tregs into three functionally distinct subpopulations based on the level of expression of Foxp3 and CD45RA:[26] (1) CD45RA$^-$Foxp3^{high} (activated Treg, aTreg), (2) CD45RA$^+$Foxp3^{low} (resting Treg, rTreg) and (3) CD45RA$^-$Foxp3^{low} (non-suppressive Foxp3^{low} cells) cells (figure 1A). We observed that the mean per cent of aTregs was significantly greater in patients in the rGCA-TCZ group (1.3% (SD 0.9)) compared with patients in the rGCA-CS group (0.6% (SD 0.4)) (p<0.01) (figure 1B). There were no significant differences among groups in terms of rTregs and non-suppressive Foxp3^{low} cells (see online supplementary figure S2). The phenotype of cellular activation observed in Tregs derived from patients in the rGCA-TCZ group was then confirmed by measuring on all CD4$^+$Foxp3$^+$ cells the expression of CCR4 and CTLA-4, markers of the most terminally differentiated activated effector Tregs[26–30] (figure 1B). The differences between patients in the rGCA-TCZ group and patients in the rGCA-CS group in terms of the numbers of aTregs, CD4$^+$Foxp3$^+$CCR4$^+$ cells and CD4$^+$Foxp3$^+$CTLA-4$^+$ cells remained statistically significant after analyses adjusted for age and CS dose. As expected, aTregs demonstrated higher expression of CD25, CCR4 and CTLA-4 when compared with rTregs and non-suppressive Foxp3^{low} cells (figure 1C). These findings demonstrate that in patients with GCA, remission maintenance with IL-6 blockade therapy is associated with increased Treg activation.

TCZ restores the impaired proliferative capacity of Tregs

Tregs are among the most actively replicating cells within the CD4$^+$T cell compartment, and impaired Treg proliferation has been implicated in the pathogenesis of autoimmunity.[31] Thus, we investigated the proliferative capacity of Tregs in patients with GCA and HCs by measuring the expression of Ki67, a

Table 1 Characteristics of the patients with GCA and the healthy individuals at baseline

	rGCA-CS (n=19)	rGCA-TCZ (n=11)	aGCA (n=11)	p Value	Controls* (n=10)	p Value
Age, years: mean (SD)	73 (10)	69 (8)	72 (10)	0.41	59 (10)	<0.01
Sex, female: number (%)	12 (63)	9 (82)	9 (82)	0.48	4 (40)	0.07
Race, white: number (%)	17 (89)	10 (91)	11 (100)	0.78	11 (100)	1.00
Relapsing disease: number (%)	12 (63)	11 (100)	8 (73)	0.06	–	–
Biopsy-proven disease: number (%)	11 (58)	5 (45)	7 (64)	0.78	–	–
Image compatible with large vessel vasculitis: number (%)†	2 (11)	4 (36)	4 (36)	0.16	–	–
Disease duration, months: median (IQR)	25.5 (9.2; 54.1)	35.7 (32.7; 70.4)	34.9 (3.7; 60.3)	0.73	–	–
Duration of CS treatment, months: median (IQR)	18.4 (9.2; 54.1)	28.4 (9.9; 67.9)	34.5 (1.0; 58.0)	0.90	–	–
Duration of TCZ treatment, months: median (IQR)	–	18 (14.2; 28.5)	–	–	–	–
Prior MTX use: number (%)	6 (32)	4 (36)	3 (27)	1.00	–	–
CS dose at time of sampling, mg/day: mean (SD)	15.7 (18.3)	0.2 (0.4)	8.0 (6.8)	0.02‡	–	–

Analysis: Analysis of variance, Kruskal-Wallis, Student's t-test and Fisher's exact test.
*Controls versus all patients with GCA.
†Indicates MR angiography, CT angiography or positron emission tomography.
‡rGCA-CS versus rGCA-TCZ.
aGCA, active GCA; CS, corticosteroids (prednisone); GCA, giant cell arteritis; MTX, methotrexate; rGCA-CS, GCA in remission on CS; rGCA-TCZ, GCA in remission on TCZ without or without CS; TCZ, tocilizumab.

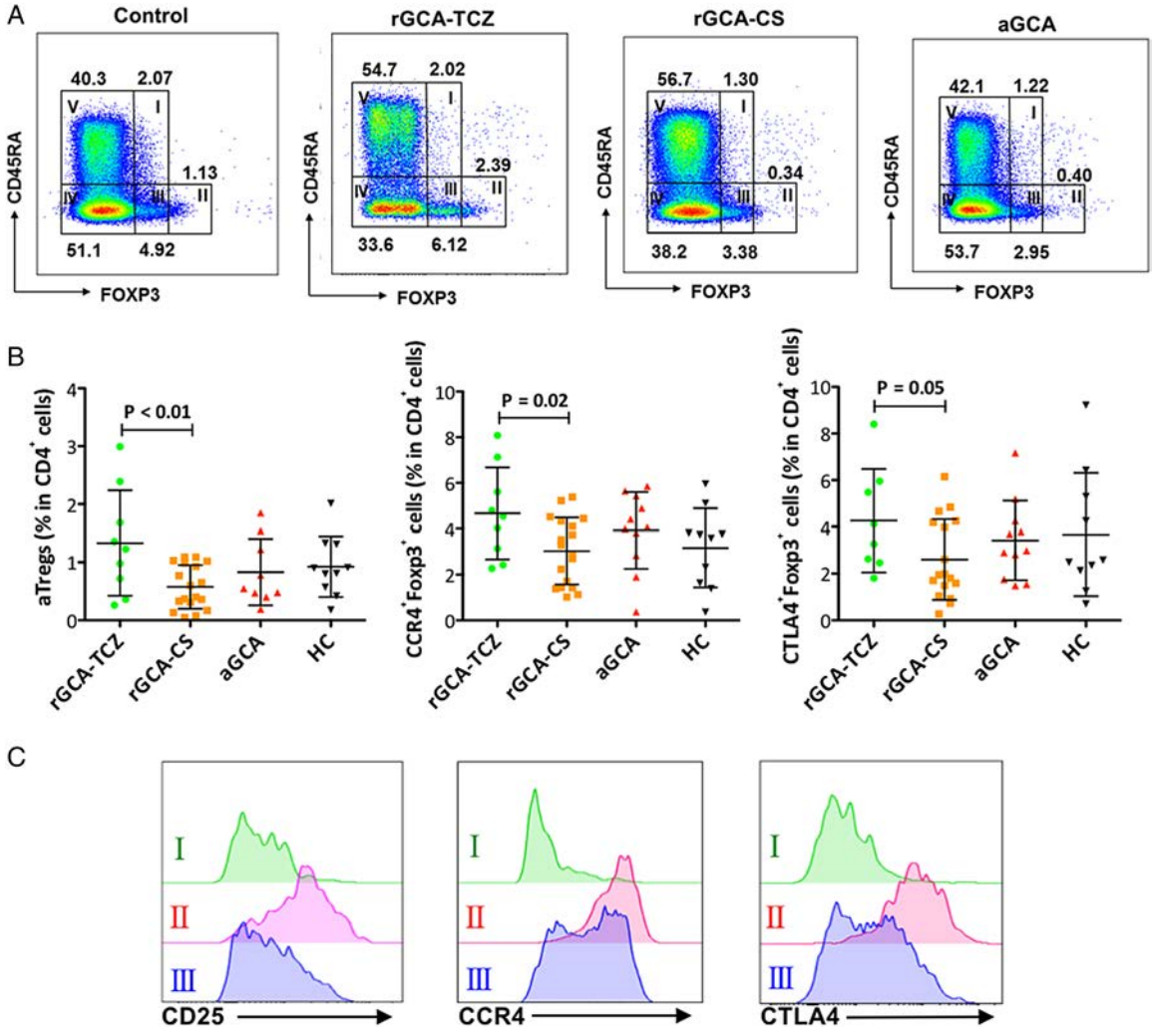


Figure 1 TCZ therapy increases the numbers of activated regulatory T (Treg) cells. $CD4^+$ T cells were purified from peripheral blood of patients with GCA and healthy controls (HCs) by negative selection. (A) Representative flow cytometry plots of $CD4^+$ T cells classified according to the expression of CD45RA and Foxp3 in (1) resting Tregs (rTregs, subset I), (2) activated Tregs (aTregs, subset II) and (3) non-suppressive $Foxp3^{low}$ cells (subset III). (B) Frequencies of aTregs, $CD4^+Foxp3^+CCR4^+$ cells and $CD4^+Foxp3^+CTLA\text{-}4^+$ cells in patients with GCA (rGCA-TCZ, n=9; rGCA-CS, n=18; aGCA, n=11) and HCs (n=10). (C) Representative histograms showing the expression of CD25, CCR4 and CTLA-4 in rTregs (I), aTregs (II) and non-suppressive $Foxp3^{low}$ cells (III). Analysis: Student's t-test. Error bars represent means and SD. aGCA, active GCA; CS, corticosteroids; GCA, giant cell arteritis; rGCA-CS, GCA in remission on CS; rGCA-TCZ, GCA in remission on TCZ without or without CS; TCZ, tocilizumab.

marker of cellular replication. We observed that the mean per cent of Ki67$^+$Tregs was equivalent in patients in the rGCA-TCZ group (27.7% (SD 9.7)) and HCs (30.0% (SD 8.6)) (p=0.71). In contrast, patients in the rGCA-TCZ group demonstrated significantly higher numbers of Ki67$^+$Tregs when compared with both, patients in the rGCA-CS group (15.4% (SD 6.4); p=0.02) and patients in the aGCA group (16.8% (SD 3.5); p=0.04) (figure 2A, B). These differences persisted in CS dose- and age-adjusted analyses. The expression of Ki67 in CD4$^+$Foxp3$^-$T cells, however, did not differ significantly among groups (see online supplementary figure S3). These results suggest that Treg proliferation is impaired in GCA and that TCZ, in contrast to CS, selectively restores the Treg replicative potential without influencing the proliferation of non-regulatory CD4$^+$T cells.

TCZ decreases the number of Foxp3Δ2$^+$ Tregs

During Treg ontogeny, the exon 2 of Foxp3 directly inhibits key transcription factors that drive the Th17 cell differentiation programme.[22 32] Recently, less suppressive Tregs that disproportionately express Foxp3Δ2 have been reported in human autoimmune disease.[24] We therefore analysed the expression of full-length Foxp3 and Foxp3Δ2 in proliferating Tregs of patients with GCA and HCs (figure 3A). We observed that the mean per cent of Ki67$^+$Foxp3Δ2$^+$Tregs was not significantly different between patients in the rGCA-TCZ group (26.3% (SD 11.4)) and HCs (25.3% (SD 6.7)) (p=0.88) (figure 3B). In contrast, patients in the rGCA-TCZ group demonstrated significantly lower numbers of Ki67$^+$Foxp3Δ2$^+$ Tregs when compared with both, patients in the rGCA-CS group (54.9% (SD 21.4); p=0.03) and patients in the aGCA group (63.2% (SD 16.2); p=0.01) (figure 3B). These differences persisted in CS dose- and age-adjusted analyses. These results demonstrate that the increased Foxp3Δ2 Treg expression seen in patients with GCA is not corrected by CS, but is abrogated upon IL-6 signalling inhibition with TCZ.

Foxp3Δ2$^+$ Tregs coexpress CD161

It has been demonstrated that IL-6-stimulated Tregs may become IL-17 producing cells.[16 20 21] In addition, IL-17-producing Tregs may also express other Th17-related markers such as CD161.[20] For this reason, we further characterised the phenotype of proliferating Tregs that either expressed full-length Foxp3 or Foxp3Δ2 by measuring the coexpression of CD25 and CD161 (figure 3C). We found that Tregs that expressed full-length Foxp3 coexpressed high amounts of CD25 (ie, CD25^{high}) significantly more often than did Tregs expressing Foxp3Δ2 (mean 74.88% (SD 13.26) vs 39.63% (SD 18.39); p<0.01). In contrast, Foxp3Δ2-expressing Tregs coexpressed CD161 significantly more often than did full-length Foxp3-expressing Tregs (mean 7.54% (SD 7.00) vs 0.66% (1.17); p<0.01) (figure 3D). The correlation between full-length Foxp3 and CD25 and between Foxp3Δ2 and CD161 was equivalent across all groups (aGCA, rGCA-CS, rGCA-TCZ and HC) (see online supplementary figure S4). In summary, Foxp3Δ2-expressing Tregs demonstrated decreased coexpression of CD25 and increased coexpression of CD161, a phenotype that suggests the potential for IL-17 production.

TCZ reduces the population of IL-17-producing Tregs

Because our data showed that proliferating Tregs derived from patients with active disease preferentially expressed Foxp3Δ2, and these cells were also characterised by CD161 co-staining, we examined the Treg production of IL-17 (figure 4A). We found that the mean per cent (SD) of IL-17$^+$Tregs in aGCA, rGCA-CS, rGCA-TCZ and HC was 4.40% (1.29), 2.68% (1.36), 1.29% (1.69) and 1.94% (1.14), respectively (figure 4B). Whereas no significant differences in the numbers of IL-17$^+$Tregs existed between HCs and patients in the rGCA-TCZ group (p=0.39), patients in the rGCA-TCZ group demonstrated significantly lower numbers of IL-17$^+$Tregs than patients in the aGCA group (p<0.01) and a trend towards fewer of these cells compared with patients in the rGCA-CS group (p=0.06). The differences among GCA groups persisted in CS dose- and age-adjusted analyses. In concordance with prior reports,[20 26] the main source of IL-17 within the Treg population of patients with active disease resided in the CD45RA$^-$Foxp3^{low} non-suppressive cell subset (figure 4C, D). These results demonstrate that the IL-17-producing Treg

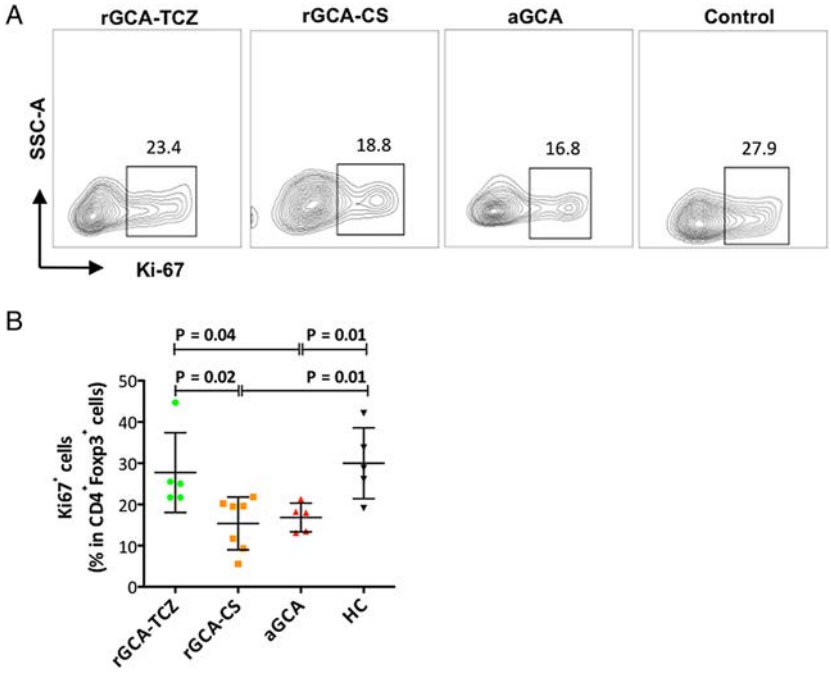


Figure 2 TCZ therapy restores impaired regulatory T (Treg) cell proliferation. (A) Representative flow cytometry plots of Ki67$^+$ cells within CD4$^+$Foxp3$^+$ T cells in patients with GCA and healthy controls (HCs). (B) Frequencies of Ki67$^+$ cells within CD4$^+$Foxp3$^+$ T cells in patients with GCA (rGCA-TCZ, n=5; rGCA-CS, n=7; aGCA, n=5) and HC (n=5). Analysis: Student's t-test. Error bars represent means and SD. aGCA, active GCA; GCA, giant cell arteritis; rGCA-CS, GCA in remission on CS; rGCA-TCZ, GCA in remission on TCZ without or without CS; TCZ, tocilizumab.

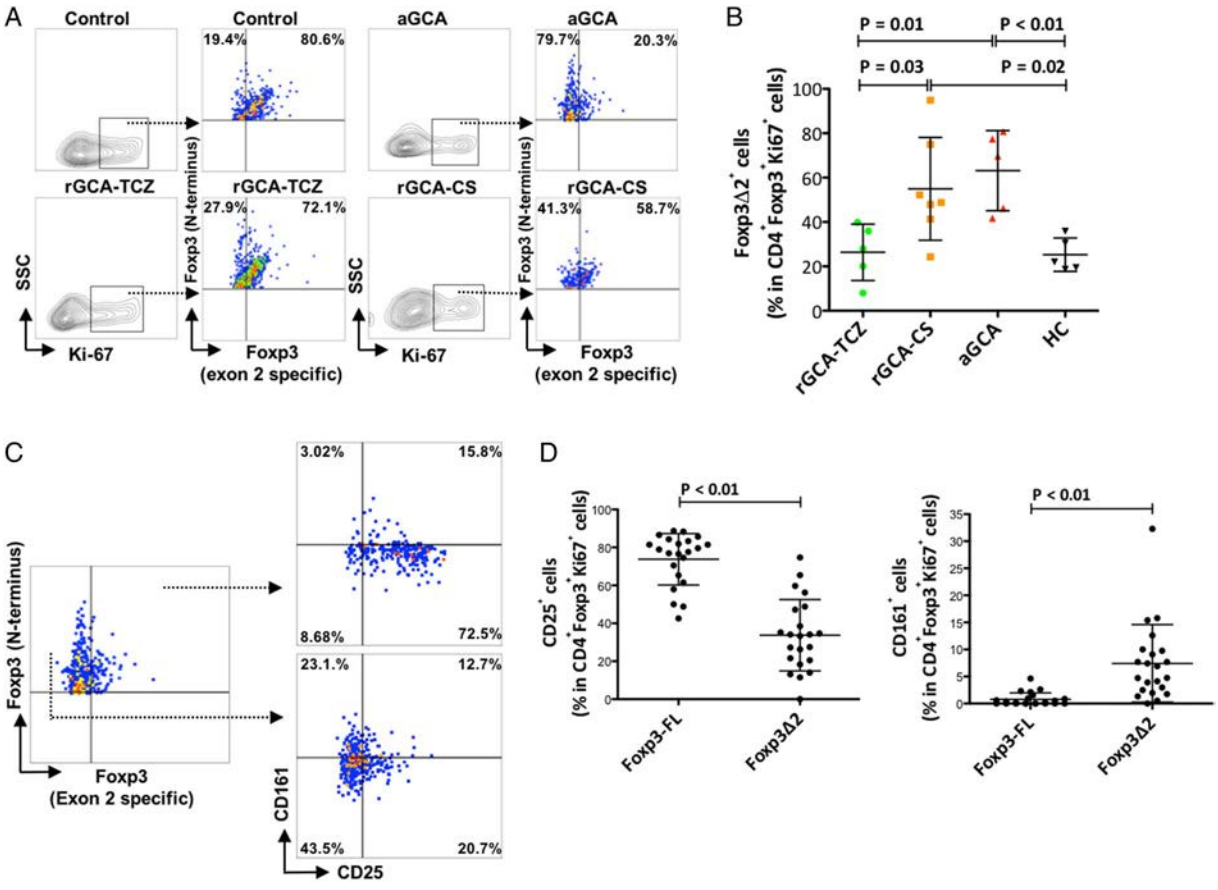


Figure 3 Effects of TCZ on Foxp3Δ2 expression in proliferating regulatory T (Treg) cells and phenotype of Foxp3Δ2$^+$ Tregs. (A) Representative flow cytometry plots of Foxp3Δ2 expression within CD4$^+$Foxp3$^+$Ki67$^+$ T cells in patients with GCA and healthy controls (HCs). (B) Frequencies of Foxp3Δ2$^+$ cells within CD4$^+$Foxp3$^+$Ki67$^+$ T cells in GCA (rGCA-TCZ, n=5; rGCA-CS, n=7; aGCA, n=5) and HCs (n=5). (C) Representative flow cytometry plots of the expression of CD25 and CD161 in full-length Foxp3-expressing and Foxp3Δ2-expressing CD4$^+$Foxp3$^+$Ki67$^+$ T cells in a patient with aGCA. (D) Surface expression of CD25 (left panel) and CD161 (right panel) in full-length Foxp3-expressing and Foxp3Δ2-expressing CD4$^+$Foxp3$^+$Ki67$^+$ T cells of patients with GCA and HCs (n=22). Analysis: unpaired Student's t-test in B; paired Student's t-test in C. Error bars represent means and SD. aGCA, active GCA; GCA, giant cell arteritis; Foxp3-FL, full-length Foxp3 isoform; rGCA-CS, GCA in remission on CS; rGCA-TCZ, GCA in remission on TCZ without or without CS; TCZ, tocilizumab.

population is expanded in peripheral blood during periods of GCA activity, and that TCZ abrogates this abnormality more efficiently than do CS.

Treg function is impaired in GCA

Previous research has shown that IL-6 may decrease Treg function.[33] In patients with new-onset GCA, however, Tregs have been reported to be competent regardless of disease activity or CS treatment.[8] To investigate whether patients with GCA in our cohort had normal or impaired Treg function and to examine whether IL-6 blockade influenced this function, we performed suppression assays coculturing CD4$^+$CD25$^-$ conventional T cells and CD4$^+$CD25$^+$ Tregs. The results showed no significant differences in Treg function among the GCA groups (figure 5A, B). However, Tregs derived from patients with GCA taken together demonstrated significantly impaired suppressive ability when compared with Tregs derived from HCs (figure 5A, B). To assess for the potential confounding effect of proinflammatory CD25$^+$ effector cells that could have been included in the population of CD4$^+$CD25$^+$ cells used for functional assays, we analysed the number of non-Tregs (CD4$^+$CD25$^+$CD45RA$^-$ cells) and non-suppressing Foxp3^{low} cells (CD4$^+$CD25^{++}CD45RA$^-$cells) in comparison to the number of aTregs (CD4$^+$CD25^{+++} CD45RA$^-$cells) and rTregs (CD4$^+$CD25^{++}CD45RA$^+$ cells) within the CD4$^+$CD25$^+$ pool[26] and found no significant differences among groups (see online supplementary figure S5).

DISCUSSION

We sought to characterise the peripheral Treg compartment in GCA and to evaluate whether IL-6R blockade was associated with modulation of the Treg response. Our results showed that patients with active disease have a defective and likely pathogenic Treg population that demonstrates decreased proliferation, overexpression of Foxp3Δ2 and increased IL-17 production. In addition, our study revealed a mechanism by which IL-6 signalling inhibition may exert its therapeutic effects in GCA.[14] Unlike therapy with low to moderate doses of CS, treatment with TCZ restored the Treg proliferative capacity, reverted the pathogenic Treg phenotype (Foxp3Δ2 and IL-17 expression) and increased the expression of markers of Treg activation, trafficking and terminal differentiation (Foxp3^{high}, CD25^{high}, CCR4 and CTLA-4).

Foxp3 largely controls the phenotype and function of Tregs.[34] Three variants of Foxp3 have been described, a full-length and two spliced forms (Foxp3Δ2 and Foxp3 lacking exon 2 and 7 (Foxp3Δ2Δ7)).[35 36] Although the regulation and function of Foxp3Δ2 and Foxp3Δ2Δ7 are poorly understood,[35 36] exon 2 is known to encode a repression domain that blocks the

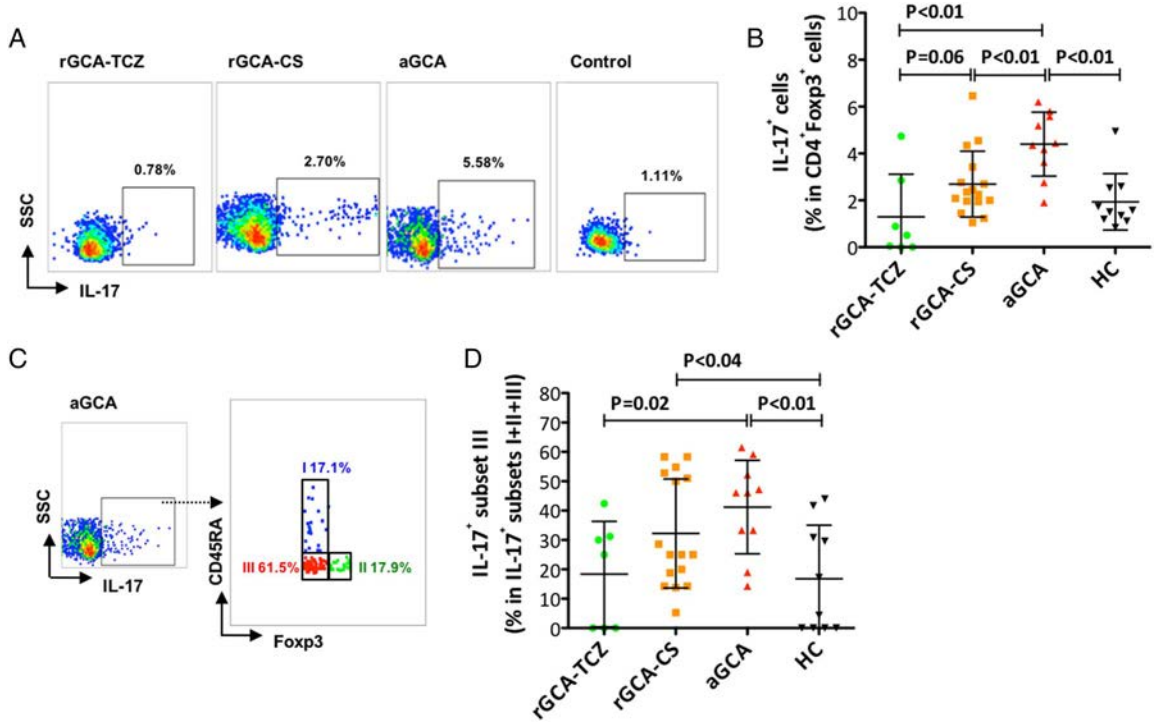


Figure 4 TCZ corrects the expansion of IL-17-producing regulatory T (Treg) cells. (A) Representative flow cytometry plots of IL-17$^+$ cells within CD4$^+$Foxp3$^+$ T cells in patients with GCA and healthy controls (HCs). (B) Frequencies of CD4$^+$Foxp3$^+$IL-17$^+$ T cells in patients with GCA (rGCA-TCZ, n=7; rGCA-CS, n=16; aGCA, n=10) and HCs (n=10). (C) Representative flow cytometry plots of IL-17 expression within rTregs (subset I), aTregs (subset II), and non-suppressive Foxp3^{low} cells (subset III) in a patient with aGCA. (D) Frequencies of IL-17 expressing non-suppressive Foxp3^{low} cells (subset III) within CD4$^+$Foxp3$^+$IL-17$^+$ T cells (subsets I+II+III) in patients with GCA (rGCA-TCZ, n=7; rGCA-CS, n=16; aGCA, n=10) and HCs (n=10). Analysis: Student's t-test. Error bars represent means and SD. aGCA, active GCA; GCA, giant cell arteritis; rGCA-CS, GCA in remission on CS; rGCA-TCZ, GCA in remission on TCZ without or without CS; TCZ, tocilizumab.

activity of the transcription factors RORγt (RORC in humans) and RORα, which are involved in the differentiation of CD4$^+$ cells towards the Th17 phenotype.[22 32 37] Foxp3Δ2 is regarded as a hypofunctional isoform of Foxp3[23 32] and increased expression of this spliced variant has been reported as a mechanism of immune dysregulation in ANCA-associated vasculitis.[24] Herein, we demonstrate for the first time that Tregs in patients with GCA preferentially express Foxp3Δ2 over full-length Foxp3. Moreover, we show that Foxp3Δ2 Tregs are often CD161^{high}CD25^{low}, which suggests potential for IL-17 production.[20] We therefore predict that Foxp3Δ2 Tregs in GCA lose their suppressive function, and themselves become pathogenic as a source of IL-17.

The functional plasticity of Tregs is highly dependent on the surrounding microenvironment,[17 18 20] and the stability of Tregs is thought to play a role in the pathogenesis of inflammatory disorders.[20 21 38 39] IL-17-producing Tregs have been detected in inflamed tissues of patients with autoimmune conditions such as rheumatoid arthritis, in which IL-6 may induce Tregs to become IL-17$^+$ cells.[20] In some cases, IL-17-producing Tregs seem to 'relax' their suppressive function,[20] but in others, they retain full regulatory capacity.[20 40 41] Of note, IL-17$^+$Foxp3$^+$ cells have also been found infiltrating arteries of patients with GCA.[10] However, their functional characterisation, contribution to disease pathogenesis and response to treatment have not been fully elucidated. Here we show that IL-17-producing Tregs are also present in peripheral blood of patients with GCA during periods of disease activity, and that their expansion normalises following IL-6R blockade therapy. In accordance to prior reports, we found that IL-17-producing Tregs in GCA also express other markers commonly associated with the Th17 lineage (eg, CD161)[20] and reside within the CD45RA$^-$Foxp3^{low} non-suppressive cell subset.[20 26] Because we observed that TCZ led to pronounced reduction of both, Foxp3Δ2 and IL-17 expression, we speculate that by a yet undefined mechanism, IL-6 promotes the transcription of Foxp3Δ2 in Tregs with subsequent polarisation of these cells towards the Th17 phenotype. The function of IL17-producing Tregs and Foxp3Δ2 Tregs in GCA remains to be determined.

The reasons why Samson *et al*[8] found reduced frequencies of functional Tregs in patients with GCA and we did not are not apparent. Possible explanations include differences in the methods used to isolate Tregs as well as differences in the characteristics of the populations analysed. Our cohort was composed of patients with GCA with long disease duration and prolonged CS exposure (mean 27 months). In contrast, the cohort studied by Samson *et al* was comprised of newly diagnosed patients, whose CS treatment was relatively short (mean 3.4 months). It is possible that early phases of the disease are characterised by decreased numbers of competent Tregs, which tend to normalise in number over time, but become functionally deficient under the influence of chronic inflammatory stimuli or prolonged CS exposures.

Tregs from patients with GCA undergoing TCZ therapy did not demonstrate enhanced ability to suppress the proliferation of conventional T cells despite the increased expression of effector molecules (eg, CTLA-4). Although this apparent discrepancy could represent a type II error, other possibilities also exist. First, an augmented regulatory response can be achieved not only by increasing Treg function, but also by increasing the trafficking of Tregs to the sites of inflammation. CCR4 is

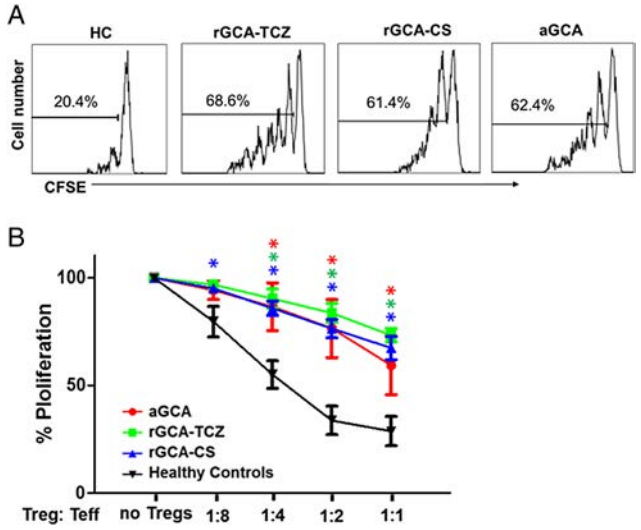


Figure 5 Regulatory T (Treg) cell function in patients with GCA and healthy controls (HCs). 10^5 CFSE-labelled $CD4^+CD25^-$ conventional T cells stimulated with anti-CD3/CD28 antibodies were incubated for 4 days with varying concentrations of autologous $CD4^+CD25^+$ Tregs (HC:HC; GCA:GCA) to create ratios of 8:1, 4:1, 2:1 and 1:1. Proliferation of conventional T cells was measured by determination of CFSE dilution by flow cytometry. (A) Conventional T cell proliferation plots from representative patients with GCA and HCs (conventional T cell to Treg ratio 1:1). (B) Suppression assays in patients with GCA (rGCA-TCZ, n=4; rGCA-CS, n=9; aGCA, n=4) and HCs (n=5). Analysis: Student's t-test. Error bars represent means and SD. Asterisks denote statistically significant differences compared with HCs. aGCA, active GCA; GCA, giant cell arteritis; rGCA-CS, GCA in remission on CS; rGCA-TCZ, GCA in remission on TCZ without or without CS; TCZ, tocilizumab; Teff=conventional T cells.

involved in Treg migration[42] and has been shown to direct Tregs into cardiovascular allografts, the allergic lung and certain tumours.[43–45] It could be hypothesised that a highly proliferative Treg compartment that expresses CCR4 may form the basis for increased Treg cell migration into inflamed arteries. In addition, other important in vivo mechanisms by which Tregs exert their function (eg, through CTLA-4 competition with CD28 for binding to CD80/86)[30 46] could not be assessed in the functional assay used.

Two randomised controlled trials have recently demonstrated that TCZ is effective in maintaining disease remission and sparing CS in GCA.[14 15] Our findings complement the results of these trials and provide a pathophysiological rationale for the use of IL-6 blockade therapy in this disease. In addition, given that several of the Treg abnormalities observed in patients with active disease were not fully reversed upon treatment with CS monotherapy, our results may also provide insight into the reason why the great majority of CS-treated patients relapse upon CS dose reduction.

In summary, we found that GCA is associated with marked abnormalities in the peripheral Treg compartment. In addition, we demonstrated that unlike CS treatment, TCZ-therapy not only abrogated the pathogenic Treg phenotype seen during periods of disease activity, but also increased Treg activation, proliferation and terminal differentiation. Limitations of our study include its cross-sectional nature and the relatively small sample size; therefore, larger studies that include longitudinal follow-up of patients with new-onset GCA treated with TCZ will be required to continue to improve our understanding of the beneficial effects of blocking IL-6 in this disorder.

Correction notice This article has been corrected since it published Epub ahead of print. The fifth sentence of the introduction has been amended and co-author details have been updated.

Contributors Substantial contributions to the conception or design of the work, or the acquisition, analysis or interpretation of data: CM, YM, KS, NDK, JHS, ADL and SU; drafting the work or revising it critically for important intellectual content: CM, YM, KS, NDK, JHS, ADL and SU; final approval of the version published: CM, YM, KS, NDK, JHS, ADL and SU; agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: ADL and SU.

Funding Arthritis Foundation CRTA grant #5924, Japan Heart Foundation and Bayer Yakuhin Research Grant Abroad.

Competing interests None declared.

Patient consent Obtained.

Ethics approval This protocol was approved by the Partners Human Research Committee. Institutional board review (IRB) protocol 2012P002348/MGH.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All data related with this study are presented in the revised manuscript and online supplementary material.

REFERENCES

1 Salvarani C, Cantini F, Hunder GG. Polymyalgia rheumatica and giant-cell arteritis. *Lancet* 2008;372:234–45.
2 Proven A, Gabriel SE, Orces C, *et al*. Glucocorticoid therapy in giant cell arteritis: duration and adverse outcomes. *Arthritis Rheum* 2003;49:703–8.
3 Jover JA, Hernández-García C, Morado IC, *et al*. Combined treatment of giant-cell arteritis with methotrexate and prednisone. A randomized, double-blind, placebo-controlled trial. *Ann Intern Med* 2001;134:106–14.
4 Hoffman GS, Cid MC, Hellmann DB, *et al*. A multicenter, randomized, double-blind, placebo-controlled trial of adjuvant methotrexate treatment for giant cell arteritis. *Arthritis Rheum* 2002;46:1309–18.
5 Hoffman GS, Cid MC, Rendt-Zagar KE, *et al*. Infliximab for maintenance of glucocorticosteroid-induced remission of giant cell arteritis: a randomized trial. *Ann Intern Med* 2007;146:621–30.
6 Seror R, Baron G, Hachulla E, *et al*. Adalimumab for steroid sparing in patients with giant-cell arteritis: results of a multicentre randomised controlled trial. *Ann Rheum Dis* 2014;73:2074–81.
7 Deng J, Younge BR, Olshen RA, *et al*. Th17 and Th1 T-cell responses in giant cell arteritis. *Circulation* 2010;121:906–15.
8 Samson M, Audia S, Fraszczak J, *et al*. Th1 and Th17 lymphocytes expressing CD161 are implicated in giant cell arteritis and polymyalgia rheumatica pathogenesis. *Arthritis Rheum* 2012;64:3788–98.
9 Terrier B, Geri G, Chaara W, *et al*. Interleukin-21 modulates Th1 and Th17 responses in giant cell arteritis. *Arthritis Rheum* 2012;64:2001–11.
10 Espígol-Frigolé G, Corbera-Bellalta M, Planas-Rigol E, *et al*. Increased IL-17A expression in temporal artery lesions is a predictor of sustained response to glucocorticoid treatment in patients with giant-cell arteritis. *Ann Rheum Dis* 2013;72:1481–7.
11 Weyand CM, Hicok KC, Hunder GG, *et al*. Tissue cytokine patterns in patients with polymyalgia rheumatica and giant cell arteritis. *Ann Intern Med* 1994;121:484–91.
12 Hernandez-Rodriguez J, Segarra M, Vilardell C, *et al*. Tissue production of pro-inflammatory cytokines (IL-1beta, TNFalpha and IL-6) correlates with the intensity of the systemic inflammatory response and with corticosteroid requirements in giant-cell arteritis. *Rheumatology (Oxf)* 2004;43:294–301.
13 Dasgupta B, Panayi GS. Interleukin-6 in serum of patients with polymyalgia rheumatica and giant cell arteritis. *Br J Rheumatol* 1990;29:456–8.
14 Villiger PM, Adler S, Kuchen S, *et al*. Tocilizumab for induction and maintenance of remission in giant cell arteritis: a phase 2, randomised, double-blind, placebo-controlled trial. *Lancet* 2016;387:1921–7.
15 Stone JH, Tuckwell K, Dimonaco S, *et al*. Efficacy and safety of tocilizumab in patients with giant cell arteritis: week 52 results from a phase 3, multicenter, randomized, double-blind, placebo-controlled trial (abstract). *American College of Rheumatology Annual Meeting Washington, DC* 2016.
16 Kleinewietfeld M, Hafler DA. The plasticity of human Treg and Th17 cells and its role in autoimmunity. *Semin Immunol* 2013;25:305–12.
17 DuPage M, Bluestone JA. Harnessing the plasticity of CD4(+) T cells to treat immune-mediated disease. *Nat Rev Immunol* 2016;16:149–63.
18 Bettelli E, Carrier Y, Gao W, *et al*. Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. *Nature* 2006;441:235–8.
19 Blatner NR, Mulcahy MF, Dennis KL, *et al*. Expression of RORgammat marks a pathogenic regulatory T cell subset in human colon cancer. *Sci Transl Med* 2012;4:164ra159.

20 Wang T, Sun X, Zhao J, *et al*. Regulatory T cells in rheumatoid arthritis showed increased plasticity toward Th17 but retained suppressive function in peripheral blood. *Ann Rheum Dis* 2015;74:1293–301.
21 Komatsu N, Okamoto K, Sawa S, *et al*. Pathogenic conversion of Foxp3+ T cells into TH17 cells in autoimmune arthritis. *Nat Med* 2014;20:62–8.
22 Ichiyama K, Yoshida H, Wakabayashi Y, *et al*. Foxp3 inhibits RORgammat-mediated IL-17A mRNA transcription through direct interaction with RORgammat. *J Biol Chem* 2008;283:17003–8.
23 Joly AL, Liu S, Dahlberg CI, *et al*. Foxp3 lacking exons 2 and 7 is unable to confer suppressive ability to regulatory T cells in vivo. *J Autoimmun* 2015;63:23–30.
24 Free ME, Bunch DO, McGregor JA, *et al*. Patients with antineutrophil cytoplasmic antibody-associated vasculitis have defective Treg cell function exacerbated by the presence of a suppression-resistant effector cell population. *Arthritis Rheum* 2013;65:1922–33.
25 Lord JD, Valliant-Saunders K, Hahn H, *et al*. Paradoxically increased FOXP3+ T cells in IBD do not preferentially express the isoform of FOXP3 lacking exon 2. *Dig Dis Sci* 2012;57:2846–55.
26 Miyara M, Yoshioka Y, Kitoh A, *et al*. Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor. *Immunity* 2009;30:899–911.
27 Sugiyama D, Nishikawa H, Maeda Y, *et al*. Anti-CCR4 mAb selectively depletes effector-type FoxP3+CD4+ regulatory T cells, evoking antitumor immune responses in humans. *Proc Natl Acad Sci USA* 2013;110:17945–50.
28 Yuan Q, Bromley SK, Means TK, *et al*. CCR4-dependent regulatory T cell function in inflammatory bowel disease. *J Exp Med* 2007;204:1327–34.
29 Read S, Malmström V, Powrie F. Cytotoxic T lymphocyte–associated antigen 4 plays an essential role in the function of CD25(+)CD4(+) regulatory cells that control intestinal inflammation. *J Exp Med* 2000;192:295–302.
30 Takahashi T, Tagami T, Yamazaki S, *et al*. Immunologic self-tolerance maintained by CD25(+)CD4(+) regulatory T cells constitutively expressing cytotoxic T lymphocyte–associated antigen 4. *J Exp Med* 2000;192:303–10.
31 Carbone F, De Rosa V, Carrieri PB, *et al*. Regulatory T cell proliferative potential is impaired in human autoimmune disease. *Nat Med* 2014;20:69–74.
32 Du J, Huang C, Zhou B, *et al*. Isoform-specific inhibition of ROR alpha-mediated transcriptional activation by human FOXP3. *J Immunol* 2008;180:4785–92.
33 Doganci A, Eigenbrod T, Krug N, *et al*. The IL-6R alpha chain controls lung CD4 +CD25+ Treg development and function during allergic airway inflammation in vivo. *J Clin Invest* 2005;115:313–25.
34 Yagi H, Nomura T, Nakamura K, *et al*. Crucial role of FOXP3 in the development and function of human CD25+CD4+ regulatory T cells. *Int Immunol* 2004;16:1643–56.
35 Allan SE, Passerini L, Bacchetta R, *et al*. The role of 2 FOXP3 isoforms in the generation of human CD4+ Tregs. *J Clin Invest* 2005;115:3276–84.
36 Smith EL, Finney HM, Nesbitt AM, *et al*. Splice variants of human FOXP3 are functional inhibitors of human CD4+ T-cell activation. *Immunology* 2006;119:203–11.
37 Zhou L, Lopes JE, Chong MM, *et al*. TGF-beta-induced Foxp3 inhibits T(H)17 cell differentiation by antagonizing RORgammat function. *Nature* 2008;453:236–40.
38 Yang XO, Nurieva R, Martinez GJ, *et al*. Molecular antagonism and plasticity of regulatory and inflammatory T cell programs. *Immunity* 2008;29:44–56.
39 Valmori D, Raffin C, Raimbaud I, *et al*. Human RORgammat+ TH17 cells preferentially differentiate from naive FOXP3+Treg in the presence of lineage-specific polarizing factors. *Proc Natl Acad Sci USA* 2010;107:19402–7.
40 Hovhannisyan Z, Treatman J, Littman DR, *et al*. Characterization of interleukin-17-producing regulatory T cells in inflamed intestinal mucosa from patients with inflammatory bowel diseases. *Gastroenterology* 2011;140:957–65.
41 Beriou G, Costantino CM, Ashley CW, *et al*. IL-17-producing human peripheral regulatory T cells retain suppressive function. *Blood* 2009;113:4240–9.
42 Iellem A, Mariani M, Lang R, *et al*. Unique chemotactic response profile and specific expression of chemokine receptors CCR4 and CCR8 by CD4(+)CD25(+) regulatory T cells. *J Exp Med* 2001;194:847–53.
43 Curiel TJ, Coukos G, Zou L, *et al*. Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nat Med* 2004;10:942–9.
44 Lee I, Wang L, Wells AD, *et al*. Recruitment of Foxp3+ T regulatory cells mediating allograft tolerance depends on the CCR4 chemokine receptor. *J Exp Med* 2005;201:1037–44.
45 Faustino L, da Fonseca DM, Takenaka MC, *et al*. Regulatory T cells migrate to airways via CCR4 and attenuate the severity of airway allergic inflammation. *J Immunol* 2013;190:2614–21.
46 Walunas TL, Lenschow DJ, Bakker CY, *et al*. CTLA-4 can function as a negative regulator of T cell activation. *Immunity* 1994;1:405–13.

OPEN ACCESS

EXTENDED REPORT

Genetic architecture distinguishes systemic juvenile idiopathic arthritis from other forms of juvenile idiopathic arthritis: clinical and therapeutic implications

Michael J Ombrello,[1] Victoria L Arthur,[1] Elaine F Remmers,[2] Anne Hinks,[3] Ioanna Tachmazidou,[4] Alexei A Grom,[5,6] Dirk Foell,[7] Alberto Martini,[8,9] Marco Gattorno,[9] Seza Özen,[10] Sampath Prahalad,[11,12] Andrew S Zeft,[13] John F Bohnsack,[14] Norman T Ilowite,[15] Elizabeth D Mellins,[16] Ricardo Russo,[17] Claudio Len,[18] Maria Odete E Hilario,[18] Sheila Oliveira,[19] Rae S M Yeung,[20,21,22] Alan M Rosenberg,[23] Lucy R Wedderburn,[24,25] Jordi Anton,[26] Johannes-Peter Haas,[27] Angela Rosen-Wolff,[28] Kirsten Minden,[29,30] Klaus Tenbrock,[31] Erkan Demirkaya,[10] Joanna Cobb,[3,32] Elizabeth Baskin,[1] Sara Signa,[8] Emily Shuldiner,[1] Richard H Duerr,[33,34] Jean-Paul Achkar,[35,36] M Ilyas Kamboh,[34] Kenneth M Kaufman,[5,6] Leah C Kottyan,[5,6] Dalila Pinto,[37] Stephen W Scherer,[38] Marta E Alarcón-Riquelme,[39,40] Elisa Docampo,[41,42] Xavier Estivill,[42,43] Ahmet Gül,[44] British Society of Pediatric and Adolescent Rheumatology (BSPAR) Study Group, Inception Cohort of Newly Diagnosed Patients with Juvenile Idiopathic Arthritis (ICON-JIA) Study Group, Childhood Arthritis Prospective Study (CAPS) Group, Randomized Placebo Phase Study of Rilonacept in sJIA (RAPPORT) Investigators, Sparks-Childhood Arthritis Response to Medication Study (CHARMS) Group, Biologically Based Outcome Predictors in JIA (BBOP) Group Carl D Langefeld,[45] Susan Thompson,[5,6] Eleftheria Zeggini,[4] Daniel L Kastner,[2] Patricia Woo,[25] Wendy Thomson[3,32]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210324).

For numbered affiliations see end of article.

Correspondence to
Dr Michael J Ombrello, Translational Genetics and Genomics Unit, Office of the Clinical Director, Intramural Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, US Department of Health and Human Services, 10 Center Drive, 12N248A, Building 10, MSC1560, Bethesda, MD 20852, USA; Michael.Ombrello@nih.gov

Received 5 August 2016
Revised 27 October 2016
Accepted 12 November 2016
Published Online First 7 December 2016

To cite: Ombrello MJ, Arthur VL, Remmers EF, *et al*. *Ann Rheum Dis* 2017;**76**:906–913.

ABSTRACT

Objectives Juvenile idiopathic arthritis (JIA) is a heterogeneous group of conditions unified by the presence of chronic childhood arthritis without an identifiable cause. Systemic JIA (sJIA) is a rare form of JIA characterised by systemic inflammation. sJIA is distinguished from other forms of JIA by unique clinical features and treatment responses that are similar to autoinflammatory diseases. However, approximately half of children with sJIA develop destructive, long-standing arthritis that appears similar to other forms of JIA. Using genomic approaches, we sought to gain novel insights into the pathophysiology of sJIA and its relationship with other forms of JIA.

Methods We performed a genome-wide association study of 770 children with sJIA collected in nine countries by the International Childhood Arthritis Genetics Consortium. Single nucleotide polymorphisms were tested for association with sJIA. Weighted genetic risk scores were used to compare the genetic architecture of sJIA with other JIA subtypes.

Results The major histocompatibility complex locus and a locus on chromosome 1 each showed association with sJIA exceeding the threshold for genome-wide significance, while 23 other novel loci were suggestive of association with sJIA. Using a combination of genetic and statistical approaches, we found no evidence of shared genetic architecture between sJIA and other common JIA subtypes.

Conclusions The lack of shared genetic risk factors between sJIA and other JIA subtypes supports the hypothesis that sJIA is a unique disease process and argues for a different classification framework. Research to improve sJIA therapy should target its unique genetics and specific pathophysiological pathways.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) encompasses a heterogeneous group of chronic childhood arthritides that develop without identifiable cause and last more than 6 weeks.[1 2] Children with JIA are placed into seven mutually exclusive categories based on clinical presentation: oligoarticular arthritis (oligoJIA) affects four or fewer joints; rheumatoid factor (RF)-negative polyarthritis (RF–polyJIA) involves five or more joints; RF-positive polyarthritis (RF+polyJIA) is analogous to adult rheumatoid arthritis; psoriatic arthritis (PsA) is an arthritis that accompanies psoriasis; enthesitis-related arthritis encompasses non-PsA childhood spondyloarthropathy; systemic arthritis (sJIA, previously known as Still's disease) is characterised by prominent systemic inflammation and has a rare adult-onset

counterpart;[3] and undifferentiated arthritis includes arthritis that does not fit into any single category.[1 2]

sJIA is among the most severe childhood inflammatory diseases. First described by Sir George Frederic Still over a century ago, sJIA is marked by arthritis and systemic inflammation with quotidian fever, evanescent salmon pink skin rash, lymphadenopathy, hepatosplenomegaly and serositis.[2 4] It is frequently complicated by macrophage activation syndrome, a potentially lethal form of hemophagocytic lymphohistiocytosis.[5] Although sJIA only constitutes approximately 10% of JIA in populations of European descent,[1 5] its disproportionately large share of the morbidity and mortality observed in JIA[6] underscores the importance of understanding and targeting its root causes.

The unique clinical characteristics of sJIA suggest that it is distinct from other forms of JIA, leading to the contention by some that sJIA should be separated from other forms of JIA and labelled as an autoinflammatory disease.[7] This has been challenged by identification of autoantibodies in some patients with sJIA.[8] Furthermore, while the systemic inflammatory features of sJIA seem to distinguish it from other forms of JIA, most children with sJIA eventually shed these features, leaving up to half of children with a persistent form of arthritis that is similar to the oligoarticular and polyarticular forms of JIA.[5 9] Finally, significant differential effects of anticytokine agents have been observed between sJIA and other forms of JIA.[10] However, due to the highly variable therapeutic responses to each agent in sJIA, this has not concretely advanced our understanding of how sJIA mechanistically relates to other forms of JIA.

One approach to evaluate the similarity of diseases is to examine shared pathophysiology through statistical comparisons of disease-specific genetic association data.[11] For example, studies of inflammatory bowel disease and spondyloarthritis have identified shared genetic risk factors, providing rationale for similar treatment choices.[11] In JIA, the majority of genetic and genomic investigations have focused on the combination of the most common subtypes, oligoJIA and RF–polyJIA (henceforth referred to in this manuscript as polygoJIA),[12 13] but until recently,[14] because of insufficient numbers of patients with sJIA, there have been only underpowered genetic studies and no genome-wide studies of sJIA. Comparisons of the genomic underpinnings of sJIA relative to other forms of JIA have therefore also been lacking.

To gain insight into the pathogenesis of sJIA, we established the International Childhood Arthritis Genetics (INCHARGE) consortium. Together, we gathered the largest sJIA study population ever assembled, which included 982 children from nine countries on three continents. Using this collection, we performed the first genome-wide association study (GWAS) of sJIA. We recently reported the results of our intensive examination of the major histocompatibility complex (MHC) locus in this study population, which identified the class II human leucocyte antigen (HLA) region as a strong sJIA susceptibility locus.[14] Here, we report the findings of the GWAS, beyond the MHC locus. Using the GWAS results, we have performed the first direct comparison of the genetic architecture of sJIA with those of the most common forms of JIA.

METHODS

Study design and participants

Peripheral blood specimens were collected from children diagnosed with sJIA according to the International League of Associations for Rheumatology (ILAR) criteria[2] by paediatric rheumatologists at participating medical centres in nine countries (see online supplementary text and figure S1). Blood samples were also obtained from geographically matched control subjects. In addition, single nucleotide polymorphism (SNP) genotype data from geographically matched control populations were used, when available. The INCHARGE project was granted institutional review board (IRB) approval by the University of Manchester. Subjects were enrolled in accordance with all local ethics regulations, with the approval of local IRBs at each contributing medical centre, and with informed parental consent.

Genotyping, quality control and imputation

Genomic DNA was extracted from peripheral blood samples. Samples were genotyped at the National Human Genome Research Institute (Bethesda, Maryland, USA) using Human Omni1M arrays (Illumina) in accordance with the manufacturer's protocols. SNP genotype data were stratified by country of origin and rigorous quality control (QC) operations were undertaken separately in each case and control population, as previously reported.[14] Principal components analysis and multidimensional scaling were used in each geographically defined case–control collection to generate nine ancestrally matched case–control strata, as previously described.[14] Genomic control inflation factors were calculated, per stratum, as an objective metric of ancestral matching.[14] An overview of the QC parameters is shown in online supplementary figure S2, and complete details are provided in the online supplementary text and our previous publication.[14]

SNP genotypes were phased using IMPUTE2,[15] and SNP imputation was performed separately for each geographically defined stratum using IMPUTE2 software and the multiancestral 1000 Genomes Project dataset (phase III) as the reference population.[16] Genotype probabilities for common markers (case minor allele frequency ≥ 0.04) that were imputed with high quality (info scores ≥ 0.8) were included in subsequent analyses.

Statistical analysis

Association testing of genotype probabilities was performed using logistic regression in each geographically defined stratum with SNPTESTv2,[15] adjusting for gender and ancestry informative principal components. Association results were meta-analysed using GWAMA.[17] Heterogeneity was evaluated in the meta-analyses using the I^2 statistic. Weighted genetic risk scores (wGRSs) were calculated and receiver operator characteristic (ROC) curve analyses were performed according to the method of Karlson *et al*.[18] wGRSs were calculated as the sum of the risk allele counts, weighted by the natural logarithm of the OR. The wGRS for polygoJIA (polygo-wGRS) incorporated 23 independent risk alleles reported by Hinks *et al*[12] (see online supplementary table S1). The wGRS for RF+polyJIA (RF+poly-wGRS) was based on the RF+polyJIA-associated wGRS-11[19] (see online supplementary table S2). The case and control distributions of risk alleles and wGRSs were evaluated with the Wilcoxon rank-sum test. Association of wGRSs with sJIA was tested by logistic regression, adjusted for ancestry and gender. The ability of wGRSs to discriminate between sJIA and other JIA subtypes was evaluated with ROC curve analysis and calculation of the area under the curve (AUC) using R. Quantile–quantile (Q–Q) plots were generated using the sJIA association data, conditional on sets of polygoJIA-associated SNPs,[12] as previously described.[20]

RESULTS

We performed SNP genotyping of 1413 children, including 982 children with sJIA and 431 healthy children. SNP genotype data, in silico, were incorporated from five existing control

populations, including 7579 additional subjects, producing a total study population of 8992 individuals. After stringent QC, 770 patients with sJIA and 6947 control subjects were stratified into nine geographically defined and ancestrally matched case–control collections (table 1, see online supplementary text and tables S3 and S4), as previously described.[14] Because most in silico control datasets were generated using SNP genotyping platforms different from that used in our study, the final number of SNPs evaluated in strata with in silico data was reduced to the intersection of the different SNP arrays (see online supplementary text and table S4). Imputation produced sets of between 4 147 566 and 6 832 892 imputed SNPs that passed postimputation QC processes (see online supplementary text). Association results were combined by fixed-effect meta-analysis, producing meta-analytic association data for 5 600 610 SNPs (figure 1). This analysis identified two sJIA susceptibility loci with associations exceeding the threshold for genome-wide significance, adjusted for the two models tested ($p<2.5\times10^{-8}$), and 23 loci with highly suggestive evidence of association ($p<5\times10^{-6}$; table 2). With the exception of the MHC locus none of these loci have been previously implicated in sJIA risk or pathophysiology. The strongest sJIA risk locus identified by this study was the MHC locus on chromosome 6 (see online supplementary figure S3). We have recently described this association in great detail in the context of a regional association study of the MHC locus in sJIA.[14] Beyond the MHC locus, we identified a novel sJIA susceptibility locus on the short arm of chromosome 1 (1p36.32) whose association also exceeded the threshold for genome-wide significance under the additive model (figures 1 and 2). This locus includes a cluster of 14 sJIA-associated SNPs that span 20.6 kb; the peak SNP is rs72632736 ($p=2.9\times10^{-9}$; OR 2.4 (1.8, 3.3). The association peak is located 20 kb upstream of *LOC284661*, a long intergenic non-coding RNA, and 263.5 kb upstream of the nearest protein coding gene, *AJAP1*, encoding adherens junction-associated protein 1. Examination of ENCODE (Encyclopedia of Noncoding DNA Elements) data revealed that the sJIA-associated SNPs overlaid a cluster of transcription factor-binding sites (TFBS) identified by chromatin immunoprecipitation sequencing (ChIP-seq; figure 2) in a variety of cell types; however, none of the top sJIA-associated SNPs were located within the ChIP-seq TFBS.

In addition to the two loci described above, this study identified 23 novel candidate susceptibility loci (figure 1, table 2),

Table 1 Summary of SNP datasets from nine sJIA case–control collections after quality control operations

Stratum	Cases	Controls	Genotyped SNPs (filtered)	Imputed	Imputed SNPs (filtered)
USA	243	1718	476 196	18 263 974	6 189 397
UK	202	4097	440 688	18 263 701	6 255 387
Germany	115	193	682 516	18 266 121	6 391 432
Turkey	49	94	682 598	18 270 612	6 389 103
Italy	49	59	686 397	18 269 173	6 375 260
Brazil	48	62	740 509	18 263 563	6 698 947
Argentina	33	115	659 100	18 263 401	6 129 601
Canada	17	427	396 935	18 263 146	5 812 530
Spain	14	182	156 136	18 261 199	4 147 550
Total	770	6947			

sJIA, systemic juvenile idiopathic arthritis; SNP, single nucleotide polymorphism.

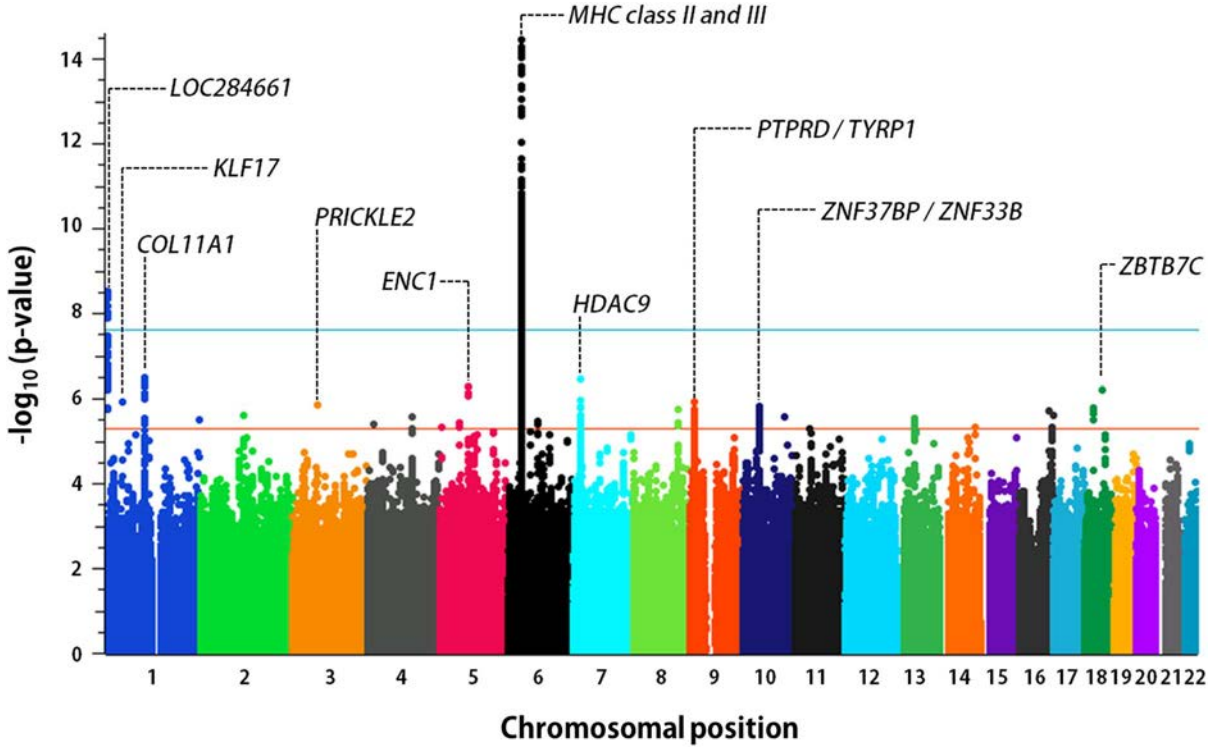


Figure 1 Genome-wide association results from meta-analysis of nine INCHARGE sJIA collections. The threshold of genome-wide significance ($p<2.5\times10^{-8}$) is shown by the blue line, while the orange line marks the level of significance suggestive of association ($p<5\times10^{-6}$). The top 10 sJIA-associated loci are labelled with the name of the nearest gene(s). INCHARGE, International Childhood Arthritis Genetics Consortium; MHC, major histocompatibility complex; sJIA, systemic juvenile idiopathic arthritis.

Table 2 Susceptibility loci with at least suggestive evidence of association with sJIA

Top SNP	Chr	Position	Ref/Alt	Best p Value	Model	OR (CI)	i^2	Strata	Samples	Closest gene(s)
rs41291794	6	32425762	A/T	3.6×10^{-15}	Additive	2.1 (1.8 to 2.6)	0.64	9	7711	*HLA-DRA*
rs72632736	1	4449204	A/G	2.9×10^{-9}	Additive	2.4 (1.8 to 3.3)	0	7	7075	*LOC284661, AJAP1*
rs1823549	1	103147831	T/C	3.2×10^{-7}	Additive	0.4 (0.3 to 0.6)	0	6	6816	*COL11A1*
rs1178121	7	18762652	C/A	3.4×10^{-7}	Dominant	1.6 (1.3 to 1.9)	0.24	8	7513	*HDAC9*
rs12517545	5	73680314	G/A	5.2×10^{-7}	Dominant	0.6 (0.5 to 0.8)	0	9	7711	*ENC1, LOC101929082*
rs79575701	18	45579621	C/A	6.2×10^{-7}	Additive	3.4 (2.1 to 5.5)	0	4	4822	*ZBTB7C*
rs114940806	1	44558672	A/G	1.2×10^{-6}	Additive	3.0 (1.9 to 4.7)	0.47	5	5137	*KLF17*
rs1279094	9	11706771	T/C	1.2×10^{-6}	Additive	1.3 (1.2 to 1.5)	0	9	7712	*LOC101929446*
rs864089	3	64244118	T/C	1.4×10^{-6}	Dominant	0.6 (0.5 to 0.8)	0	8	7516	*PRICKLE2*
rs481331	10	43003048	A/T	1.4×10^{-6}	Additive	1.5 (1.3 to 1.7)	0	9	7712	*ZNF37BP, ZNF33B*
rs8097070	18	23086307	A/G	1.6×10^{-6}	Additive	0.3 (0.2 to 0.5)	0	4	4993	*ZNF521, SS18*
rs1527934	8	117392156	C/T	1.8×10^{-6}	Additive	1.7 (1.4 to 2.1)	0	6	6926	*EIF3H, LINC00536*
rs78507369	16	78305293	A/G	2.0×10^{-6}	Additive	3.0 (1.9 to 4.6)	0	4	4857	*WWOX, LSM3P5*
rs12445022	16	87575332	G/A	2.4×10^{-6}	Dominant	1.5 (1.3 to 1.7)	0.36	9	7715	*LOC101928737, JPH3*
rs112165031	2	112902227	G/A	2.5×10^{-6}	Additive	2.5 (1.7 to 3.7)	0.58	5	6917	*FBLN7*
rs6853094	4	116576274	C/A	2.6×10^{-6}	Additive	1.6 (1.3 to 1.9)	0.22	8	7564	*RPF2P2, PGAM4P2*
rs73401585	10	109690236	T/C	2.6×10^{-6}	Additive	3.2 (2.0 to 5.2)	0	4	4824	*LOC101927573, SORCS1*
rs9595973	13	49286438	G/A	2.8×10^{-6}	Dominant	2.8 (1.8 to 4.3)	0.5	4	6845	*CYSLTR2*
rs9633402	1	247946160	G/A	3.0×10^{-6}	Dominant	0.4 (0.2 to 0.6)	0	9	7708	*TRIM58*
rs62438583	6	75326244	T/G	3.4×10^{-6}	Dominant	0.7 (0.6 to 0.8)	0	9	7712	*LOC101928516, COL12A1*
rs62359376	5	52411328	G/A	3.6×10^{-6}	Dominant	1.7 (1.4 to 2.2)	0.13	8	7516	*LOC257396, MOCS2*
rs1501138	4	16397067	T/C	4.0×10^{-6}	Dominant	0.3 (0.2 to 0.5)	0.24	8	7517	*LDB2, TAPT1, ZEB2P1*
rs7712113	5	4985443	G/C	4.5×10^{-6}	Dominant	3.7 (2.1 to 6.5)	0.68	4	4661	*LINC01020, LOC101929176*
rs1885747	14	93047455	A/G	4.6×10^{-6}	Additive	1.4 (1.2 to 1.7)	0.38	8	7513	*RIN3, LGMN*
rs111580313	16	86621219	C/T	4.8×10^{-6}	Dominant	1.7 (1.4 to 2.2)	0	7	7368	*MTHFSD, FOXL1, FOXC2*

Best p value, meta-analytic p value corrected for gender and ancestry under the model specified in the Model column. Model, the genetic model (either additive or dominant) that showed the strongest association between the SNP and sJIA. I^2, I^2 test for heterogeneity. Strata, number of strata included in meta-analysis. Samples, number of samples included in meta-analysis.
Alt, alternate allele; Chr, chromosome; Ref, reference allele; sJIA, systemic juvenile idiopathic arthritis; SNP, single nucleotide polymorphism;

two of which are shown in detail in online supplementary figure S4. Importantly, the top 25 sJIA susceptibility loci had scant intersection with the known susceptibility loci of other JIA subtypes. Based on this observation, we sought to compare the genetic architecture of sJIA with those of polygoJIA and RF+polyJIA.

We first examined the 23 polygoJIA-associated loci reported by Hinks *et al*[12] in the sJIA study population and none showed even a modest association with sJIA (see online supplementary tables S5 and S6). To more formally compare sJIA with polygoJIA, we calculated a polygo-wGRS in the sJIA case–control collections based on the same 23 SNPs. The non-parametric Wilcoxon rank-sum test found no difference in the distribution of polygoJIA risk allele counts or polygo-wGRSs between sJIA cases and controls (figure 3, see online supplementary table S7 and figures S5 and S6). Consistent with this, logistic regression analysis found no correlation between the polygo-wGRS and sJIA in any individual stratum or in the full study population (see online supplementary table S7). Analysis of ROC curves in individual strata and the full population found that the AUCs for polygo-wGRS were all close to 0.5, indicating that the polygo-wGRS was no better than random chance at distinguishing sJIA cases from control subjects (figure 3, see online supplementary figure S7). Finally, to expand the scope of our comparison beyond peak SNPs from risk loci, we performed a Q–Q plot-based enrichment analysis to look for shared genetic risk factors between sJIA and polygoJIA (figure 3). By comparing Q–Q plots of polygoJIA-associated SNPs[12] at several different significance levels in our sJIA collection, we sought to evaluate pleiotropy in a more global/genomic manner. In the presence of pleiotropy, the slopes of the Q–Q plots of disease A associations are expected to increase as the plotted SNP sets become more strongly associated with disease B, as previously shown.[20] In the case of polygoJIA-associated SNPs in sJIA, the slope of the Q–Q plots of sJIA associations did not increase when SNPs of increasingly strong association with polygoJIA were plotted, indicating that there was no enrichment of sJIA-associated variants among polygoJIA-associated variants, and therefore that there was no evidence of pleiotropy (figure 3).

In addition, we used an RF+poly-wGRS[19] to look for shared genetic architecture between sJIA and RF+polyJIA. As was the case with polygoJIA, non-parametric testing revealed no significant difference in the distribution of RF+polyJIA risk alleles (see online supplementary figure S8) or RF+poly-wGRS (see online supplementary figure S9) between sJIA cases and controls in any individual population. Of note, non-parametric testing and logistic regression analysis identified a significant difference in RF+poly-wGRS between sJIA and controls in the full collection (see online supplementary table S8 and figure S10); however, the wGRSs were actually lower in the sJIA cases than in the controls (see online supplementary figure S11). Consistent with these observations, ROC analyses found that the RF+poly-wGRS was not predictive of sJIA (see online

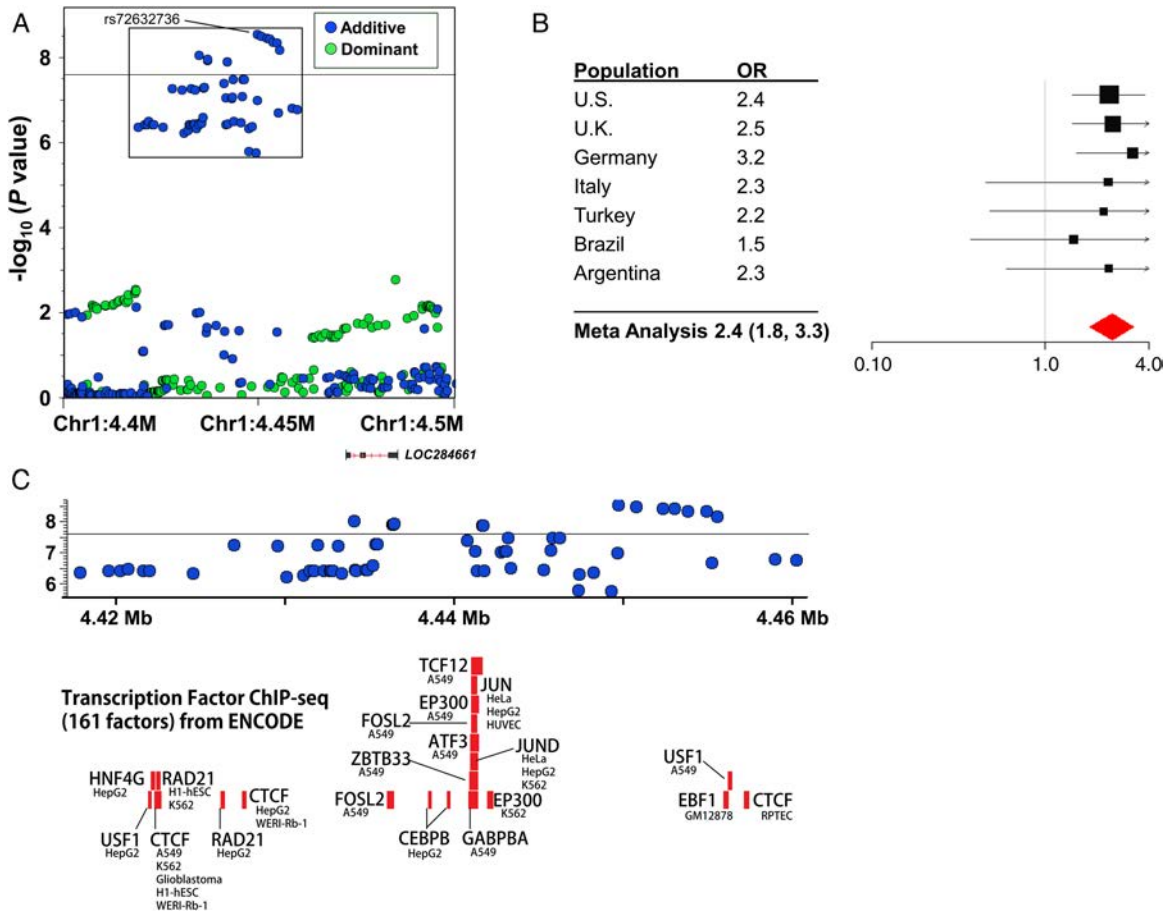


Figure 2 Systemic juvenile idiopathic arthritis (sJIA) susceptibility locus at chr1p36.32. A regional association plot demonstrates the association between sJIA) and single nucleotide polymorphisms (SNPs) in this region (A). The effect of the peak SNP (rs72632736) in each study population is demonstrated in the forest plot (B). The threshold of genome-wide significance ($p<2.5\times10^{-8}$) is marked by the black horizontal line in (A) and (C). Panel C shows the superimposition of sJIA-associated SNPs (inset box, A) with transcription factor-binding sites determined by chromatin immunoprecipitation (ChIP) sequencing from the Encyclopedia of Noncoding DNA Elements (ENCODE) project.

supplementary figures S10 and S12). Collectively, these investigations failed to identify any evidence of shared genetic architecture between sJIA and polygoJIA or RF+polyJIA.

DISCUSSION

In this study, two novel susceptibility loci met genome-wide significance criteria for association with sJIA and 23 other loci demonstrated highly suggestive evidence of association. Furthermore, formal comparisons of association data from sJIA with those from polygoJIA and RF+polyJIA have demonstrated that sJIA bears a unique genetic architecture, indicating that its underlying pathophysiological mechanisms are significantly divergent from other forms of JIA. This has important implications and should direct research for future targets of therapeutic intervention for children affected with sJIA.

This is the first large-scale genomic study of sJIA, which includes case–control collections from nine different countries. In a sample of 982 affected children, we identified genome-wide significant evidence of association with SNPs in the class II MHC locus and SNPs on chromosome 1 nearest to an uncharacterised long non-coding RNA gene. This work also identified many additional candidate sJIA susceptibility loci, nearly all of them novel, and aside from the HLA locus, none of these novel loci are associated with any other rheumatic diseases (see online supplementary table S9). The identification of these loci is an important step towards the elucidation of the specific pathways and pathogenic mechanisms in sJIA, which in turn will allow the development of therapies to more specifically target sJIA pathophysiology in affected children. Several of the susceptibility loci that warrant further investigation include strong candidates for therapeutic modulation, and many novel loci or genes that have been poorly studied, to date. Functional investigations are needed to identify and understand the specific mechanisms that underlie the genetic associations.

This study also provided the first opportunity to demonstrate that sJIA did not share heritable risk factors with the more common oligoarticular and polyarticular forms of JIA. There was no intersection of the top susceptibility loci of sJIA with those of polygoJIA or RF+polyJIA. Even within the class II MHC region, which harbours disease-associated genetic variation in each of these categories of JIA, the subtype-specific risk factors (SNPs, HLA alleles and HLA haplotypes) are not shared between subtypes. Using a combination of genetic risk scores and enrichment analysis, this study reveals an absence of shared genetic architecture between sJIA and either polygoJIA or RF+polyJIA, despite often sharing a chronic arthritis feature with polygo or RF+polyJIA. It could be that as a clinical feature, arthritis is a non-specific finding that is present in many different conditions, including infections, malignancies, autoimmune disorders and autoinflammatory conditions. These distinct genetic data provide hard evidence that these conditions differ in pathophysiology, strongly supporting the clinical distinction

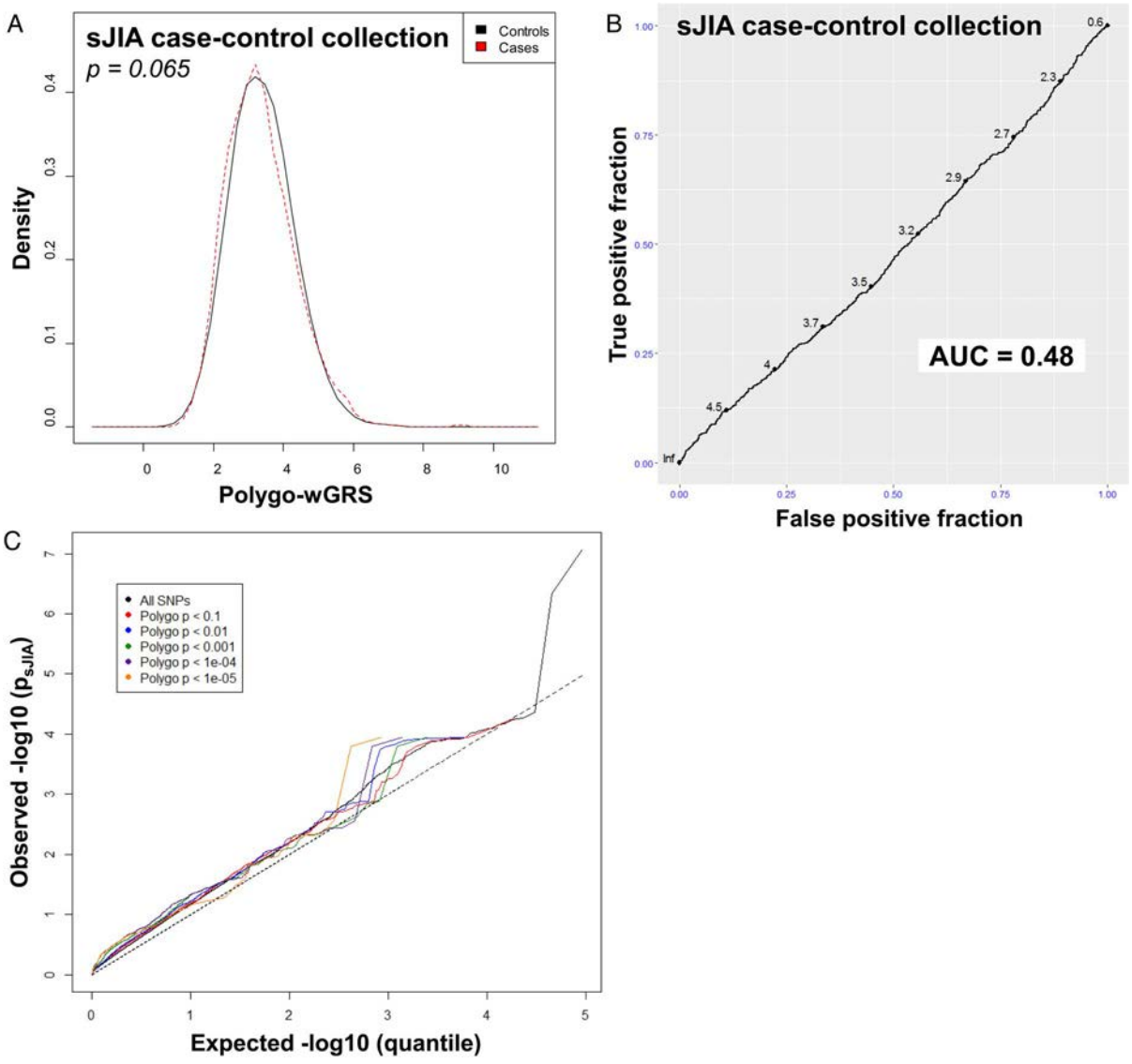


Figure 3 Comparison of the genetic architecture of systemic juvenile idiopathic arthritis (sJIA) with seronegative polyarticular and oligoarticular (polygo) JIA. Kernel density plots display the distribution of polygo-wGRS in sJIA cases and controls from the full study collection (A). p Value was calculated with the Wilcoxon rank-sum test. Receiver operator characteristic (ROC) curves with area under the curve (AUC) calculations demonstrate the performance of polygo-wGRS at predicting sJIA status in the full collection (B). Q–Q plots show the level of association of subsets of polygoJIA-associated single nucleotide polymorphisms in the sJIA population (C).

between sJIA and the other JIA subtypes. Considering the ongoing discussions about restructuring the JIA nomenclature, these studies will help inform and guide the debate surrounding sJIA[7] and how it should be classified.

The genetic dissimilarity of sJIA and other JIA subtypes has important therapeutic implications for children with sJIA. Currently, the treatment of sJIA presents physicians with a clinical conundrum, with no single, universally effective therapeutic approach. Prior to the era of biological response-modifying agents, sJIA was treated with disease-modifying antirheumatic drugs, including methotrexate, with a rationale for use extrapolated from other forms of JIA; there were no clinical trials and only limited outcome studies describing their effectiveness in sJIA.[10] In the absence of clear therapeutic alternatives, and despite the limited evidence of efficacy, methotrexate remains an accepted therapeutic option in the consensus treatment protocol.[21] Similarly, therapies targeting the cytokine tumour necrosis factor-α are highly effective in the treatment of other forms of JIA,[22] but show only modest effect in children with sJIA.[10] Today, even with the most effective treatments for sJIA directed against the inflammatory cytokines interleukin (IL)-1 and IL-6,[10] a sizable proportion of children continue to have active disease, with chronic arthritis persisting in nearly 40% of children in a recent study.[9] Currently, the only widely effective treatment for sJIA remains large doses of glucorticoids.[10] There is clearly an imperative to look for root causes of sJIA to identify better targets for therapy and prevent the development of persistent, disabling arthritis.

Although it is necessary to better understand the function of the risk alleles identified by this study, the results may identify genetic profiles that can be used to determine appropriate therapeutic interventions. To this point, two susceptibility loci are of particular therapeutic interest in sJIA: the class II HLA locus and *HDAC9*, encoding histone deacetylase 9. Given that class II HLA molecules present peptide antigens to T-cell receptors on CD4+ T cells, resulting in their activation, one may predict that

therapeutic modulation of T-cell activation would be an effective strategy in the treatment of sJIA. In fact, abatacept, which reduces T-cell activation through costimulatory inhibition, has shown promising results in children with the chronic, persistent arthritis of sJIA[23 24]—a subset of patients with sJIA who are particularly refractory to therapeutic intervention.[5] Based on these observations, it may be reasonable to use abatacept in children with sJIA. HDAC9 confers important epigenetic effects through deacetylation of histone proteins, while also regulating critical innate immune processes, including Toll-like receptor signalling and the development of regulatory T cells, via deacetylation of non-histone targets.[25–28] Despite the fact that HDAC9 was only suggestively associated with sJIA, a pilot study of the non-specific HDAC inhibitor, gavinostat, produced promising preliminary results in children with sJIA,[29] raising the possibility that HDAC inhibition represents another plausible targeted therapeutic strategy in sJIA.

At a time when an emphasis is being placed on the personalisation of medicine, it is important that we move away from broad classifications based on non-specific clinical observations and move towards the use of molecular and genetic data in establishing diagnoses, as well as pathophysiology. In turn, clinical practice will advance as these data are translated into targeted therapeutic approaches. Perhaps it is time to separate this condition from JIA all together to make clear that it is fundamentally different from any other form of JIA and needs to be considered and treated differently. Given that the currently available treatments for this condition are still imperfect, it remains imperative to continue to employ contemporary investigative approaches in sJIA, to elucidate its pathophysiology and to identify the next generation of therapeutic strategies.

Author affiliations
[1]Translational Genetics and Genomics Unit, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, US Department of Health and Human Services, Bethesda, Maryland, USA
[2]Inflammatory Disease Section, National Human Genome Research Institute, National Institutes of Health, US Department of Health and Human Services, Bethesda, Maryland, USA
[3]Arthritis Research UK Centre for Genetics and Genomics, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, Manchester, UK
[4]Human Genetics, The Wellcome Trust Sanger Institute, Hinxton, UK
[5]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA
[6]Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA
[7]Department of Pediatric Rheumatology and Immunology, University Hospital Münster, Münster, Germany
[8]Department of Pediatrics, University of Genova, Genoa, Italy
[9]Pediatrics II Unit, Giannina Gaslini Institute, Genoa, Italy
[10]Department of Pediatric Rheumatology, Hacettepe University, Ankara, Turkey
[11]Departments of Pediatrics and Human Genetics, Emory University School of Medicine, Atlanta, Georgia, USA
[12]Children's Healthcare of Atlanta, Atlanta, Georgia, USA
[13]Department of Pediatrics, Cleveland Clinic, Cleveland, Ohio, USA
[14]Department of Pediatrics, University of Utah, Salt Lake City, Utah, USA
[15]Department of Pediatrics, Albert Einstein College of Medicine and Children's Hospital at Montefiore, Bronx, New York, USA
[16]Department of Pediatrics, Stanford University, Stanford, California, USA
[17]Service of Immunology and Rheumatology, Hospital de Pediatria Garrahan, Buenos Aires, Argentina
[18]Department of Pediatrics, Universidade Federal de São Paulo, São Paulo, Brazil
[19]Universidade Federal de Rio de Janeiro, Rio de Janeiro, Brazil
[20]Department of Pediatrics, University of Toronto, Toronto, Canada
[21]Department of Immunology, University of Toronto, Toronto, Canada
[22]Institute of Medical Science, University of Toronto, Toronto, Canada
[23]Department of Pediatrics, University of Saskatchewan, Saskatoon, Canada
[24]Institute of Child Health, University College London, London, UK
[25]Center of Paediatric and Adolescent Rheumatology, University College London, London, UK
[26]Pediatric Rheumatology Unit, Hospital Sant Joan de Déu, Universitat de Barcelona, Barcelona, Spain
[27]German Center for Pediatric and Adolescent Rheumatology, Garmisch-Partenkirchen, Germany
[28]University Hospital Cal Gustav Carus, Dresden, Germany
[29]Department of Rheumatology and Clinical Immunology, Charité -University Medicine, Berlin, Germany
[30]Epidemiology Unit, German Rheumatism Research Centre, Berlin, Germany
[31]Department of Pediatrics, RWTH Aachen University, Aachen, Germany
[32]National Institute for Health Research Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK
[33]Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[34]Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[35]Department of Gastroenterology and Hepatology, Cleveland Clinic, Cleveland, Ohio, USA
[36]Department of Pathobiology, Cleveland Clinic, Cleveland, Ohio, USA
[37]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA
[38]The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada
[39]Center for Genomics and Oncological Research, Pfizer-University of Granada-Andalusian Government, Granada, Spain
[40]Unit of Chronic Inflammatory Diseases, Institute for Environmental Medicine, Karolinska Institutet, Solna, Sweden
[41]Interdisciplinary Cluster for Applied Genoproteomics-Université de Liège, Liège, Belgium
[42]Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, and Universitat Pompeu Fabra (UPF), Barcelona, Spain
[43]Sidra Medical and Research Centre, Doha, Qatar
[44]Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey
[45]Department of Biostatistical Sciences, Wake Forest University Health Sciences, Winston-Salem, North Carolina, USA

Twitter Follow NIAMS/NIH/DHHS @NIH_NIAMS; Ricardo Russo @el_reumatologo; University of Manchester, Centre for Musculoskeletal Disease @UoMMskResearch

Collaborators Full membership of collaborating consortia are listed in the supplementary text: British Society of Pediatric and Adolescent Rheumatology Study Group, Inception Cohort of Newly Diagnosed Patients with Juvenile Idiopathic Arthritis Study Group, Childhood Arthritis Prospective Study Group, Randomized Placebo Phase Study of Rilonacept in sJIA Investigators, Sparks-Childhood Arthritis Response to Medication Study Group and Biologically Based Outcome Predictors in JIA Group.

Contributors All authors participated in study design; AAG, DF, AM, MG, SÖ, SP, ASZ, JFB, NTI, EDM, RR, CL, MOEH, SO, RSMY, AMR, LRW, JA, J-PH, AR-W, KM, KT, ED, BSPAR, ICON-JIA, CAPS, RAPPORT, CHARMS, BBOP, RHD, JPA, MIK, KMK, LCK, DP, SWS, MEA-R, ED, XE and AG provided samples for the study; MJO, VLA, EFR, AH, IT, EZ, PW and WT performed the research and analysis and interpreted the data; all authors drafted and/or substantively edited the manuscript and have thoroughly reviewed and approved of the content.

Funding This work was supported by the Intramural Research Programs of the National Institute of Arthritis and Musculoskeletal and Skin Diseases (Z01-AR041198 to MJO) and the National Human Genome Research Institute (Z01-HG200370 to DLK) of the National Institutes of Health (NIH). Additional funding was provided by NIH grants R01-AR059049 (AAG), R01AR061297 (EDM), R01-AR060893 (SP), P30-AR47363 and P01-AR48929 (ST), AG030653, AG041718 and AG005133 (MIK) and U01-DK062420 and R01-DK076025 (RHD); Arthritis Research UK Grant 20385 (WT); the German Federal Ministry of Education and Research (BMBF project 01ER0813) for the 'ICON-JIA' inception cohort (KM and DF); the Val A. Browning Charitable Foundation (JFB); the Marcus Foundation (SP); the Proyecto de Excelencia (CTS-2548) of the Andalousian Government (MA-R) and the Swedish Association Against Rheumatism (MA-R). IT and EZ were supported by the Wellcome Trust (098051). WT and JC are funded by the National Institute for Health Research Biomedical Research Unit Funding Scheme. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. The CAPS study was funded by Arthritis Research UK Grant 20542. WT, AH, and JC are supported by the Manchester Academic Health Sciences Centre (MAHSC). SPARKS-CHARMS was funded by grants from SPARKS UK (08ICH09 and 12ICH08), the Medical Research Council (MR/M004600/1) and the UK National Institute for Health Research GOSH Biomedical Research Centre. The BBOP study was supported by the Canadian Institutes of Health Research and the Arthritis Society (CIHR funding reference number 82517) and the Canadian Arthritis Network (funding reference SRI-IJD-01). This research was supported in part by the Cincinnati Children's

Research Foundation and its Cincinnati Genomic Control Cohort. The authors acknowledge the use of DNA from the UK Blood Services collection of Common Controls (UKBS-CC collection), which is funded by Wellcome Trust grant 076113/C/04/Z and by the USA NIH research programme grant to the National Health Service Blood and Transplant (RP-PG-0310-1002). The authors acknowledge the use of DNA from the British 1958 Birth Cohort collection, which is funded by the UK Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. This study used the computational resources of the Biowulf system at the NIH, Bethesda, MD (http://biowulf.nih.gov).

Competing interests AAG: consulting fees and honoraria from Novimmune, Novartis, Roche. AM: consulting fees and honoraria from Abbvie, Boehringer, Celgene, CrescendoBio, Janssen, Meddimune, Novaris, NovoNordisk, Pfizer, Sanofi Aventis, Vertex and Servier, contributions have been received by G. Gaslini Hospital and the PRINTO network by BMS, GlaxoSmithKline, Hoffman-La Roche, Novartis, Pfizer, Sanofi Aventis, Schwarz Biosciences, Abbot, Francesco Angelini S.P.A., Sobi, and Merck Serono. MG: consulting and speaker fees from SOBI and Novartis, unrestricted grants to Eurofever from SOBI and Novartis. SP: consulting fees from Novartis. EDM: consulting fees from Novartis. LRW: speaker fees from Pfizer.

Ethics approval University of Manchester.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The quality control processed directly genotyped SNP genotype data from sJIA cases genotyped for this study will be deposited into the National Institutes of Health's Database of Genotypes and Phenotypes, where allowable by the Ethics and consent documents. The future use of these data will be dictated by the terms of Ethics and consent documents and the institutional certifications provided by the collaborating centres.

REFERENCES

1 Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011;377:2138–49.
2 Petty RE, Southwood TR, Manners P, *et al*. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004;31:390–2.
3 Bannatyne GA, Wohlmann AS. Rheumatoid arthritis: It's clinical history, etiology, and pathology. *Lancet* 1896;147:1120–5.
4 Still GF. On a form of chronic joint disease in children. *Med Chir Trans* 1897;80:47–60.
5 Mellins ED, Macaubas C, Grom AA. Pathogenesis of systemic juvenile idiopathic arthritis: some answers, more questions. *Nat Rev Rheumatol* 2011;7:416–26.
6 Davies R, Southwood T, Kearsley-Fleet L, *et al*. Standardized mortality rates are increased in patients with severe juvenile idiopathic arthritis. *Rheumatology* 2015;54:153.
7 Martini A. It is time to rethink juvenile idiopathic arthritis classification and nomenclature. *Ann Rheum Dis* 2012;71:1437–9.
8 Hugle B, Hinze C, Lainka E, *et al*. Development of positive antinuclear antibodies and rheumatoid factor in systemic juvenile idiopathic arthritis points toward an autoimmune phenotype later in the disease course. *Pediatr Rheumatol Online J* 2014;12:28.
9 Janow G, Schanberg LE, Setoguchi S, *et al*. The systemic juvenile idiopathic arthritis cohort of the childhood arthritis and rheumatology research alliance registry: 2010-2013. *J Rheumatol* 2016;43:1755–62.
10 Beukelman T. Treatment advances in systemic juvenile idiopathic arthritis. *F1000Prime Rep* 2014;6:21.
11 Solovieff N, Cotsapas C, Lee PH, *et al*. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* 2013;14:483–95.
12 Hinks A, Cobb J, Marion MC, *et al*. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nat Genet* 2013;45:664–9.
13 Thompson SD, Marion MC, Sudman M, *et al*. Genome-wide association analysis of juvenile idiopathic arthritis identifies a new susceptibility locus at chromosomal region 3q13. *Arthritis Rheum* 2012;64:2781–91.
14 Ombrello MJ, Remmers EF, Tachmazidou I, *et al*. HLA-DRB1*11 and variants of the MHC class II locus are strong risk factors for systemic juvenile idiopathic arthritis. *Proc Natl Acad Sci USA* 2015;112:15970–5.
15 Marchini J, Howie B, Myers S, *et al*. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007;39:906–13.
16 Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* 2011;1:457–70.
17 Magi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 2010;11:288.
18 Karlson EW, Chibnik LB, Kraft P, *et al*. Cumulative association of 22 genetic variants with seropositive rheumatoid arthritis risk. *Ann Rheum Dis* 2010;69:1077–85.
19 Prahalad S, Conneely KN, Jiang Y, *et al*. Susceptibility to childhood-onset rheumatoid arthritis: investigation of a weighted genetic risk score that integrates cumulative effects of variants at five genetic loci. *Arthritis Rheum* 2013;65:1663–7.
20 Liley J, Wallace C. A pleiotropy-informed Bayesian false discovery rate adapted to a shared control design finds new disease associations from GWAS summary statistics. *PLoS Genet* 2015;11:e1004926.
21 DeWitt EM, Kimura Y, Beukelman T, *et al*. Consensus treatment plans for new-onset systemic juvenile idiopathic arthritis. *Arthritis Care Res (Hoboken)* 2012;64:1001–10.
22 Ungar WJ, Costa V, Burnett HF, *et al*. The use of biologic response modifiers in polyarticular-course juvenile idiopathic arthritis: a systematic review. *Semin Arthritis Rheum* 2013;42:597–618.
23 Ruperto N, Lovell DJ, Quartier P, *et al*. Abatacept in children with juvenile idiopathic arthritis: a randomised, double-blind, placebo-controlled withdrawal trial. *Lancet* 2008;372:383–91.
24 Record JL, Beukelman T, Cron RQ. Combination therapy of abatacept and anakinra in children with refractory systemic juvenile idiopathic arthritis: a retrospective case series. *J Rheumatol* 2011;38:180–1.
25 Birnbaum RY, Clowney EJ, Agamy O, *et al*. Coding exons function as tissue-specific enhancers of nearby genes. *Genome Res* 2012;22:1059–68.
26 Li X, Zhang Q, Ding Y, *et al*. Methyltransferase Dnmt3a upregulates HDAC9 to deacetylate the kinase TBK1 for activation of antiviral innate immunity. *Nat Immunol* 2016;17:806–15.
27 Li B, Samanta A, Song X, *et al*. FOXP3 interactions with histone acetyltransferase and class II histone deacetylases are required for repression. *Proc Natl Acad Sci USA* 2007;104:4571–6.
28 Clark RI, Tan SW, Pean CB, *et al*. MEF2 is an in vivo immune-metabolic switch. *Cell* 2013;155:435–47.
29 Vojinovic J, Damjanov N, D'Urzo C, *et al*. Safety and efficacy of an oral histone deacetylase inhibitor in systemic-onset juvenile idiopathic arthritis. *Arthritis Rheum* 2011;63:1452–8.

EXTENDED REPORT

CCL2/CCR2, but not CCL5/CCR5, mediates monocyte recruitment, inflammation and cartilage destruction in osteoarthritis

Harini Raghu,[1,2] Christin M Lepus,[1,2] Qian Wang,[1,2] Heidi H Wong,[1,2] Nithya Lingampalli,[1,2] Francesca Oliviero,[3] Leonardo Punzi,[3] Nicholas J Giori,[2,4] Stuart B Goodman,[5] Constance R Chu,[2,4] Jeremy B Sokolove,[1,2] William H Robinson[1,2]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210426).

[1]Division of Immunology and Rheumatology, Stanford University School of Medicine, Stanford, California, USA
[2]VA Palo Alto Health Care System, Palo Alto, California, USA
[3]Rheumatology Unit, Department of Medicine, University of Padova, Padova, Italy
[4]Department of Orthopedic Surgery, Stanford University School of Medicine, Stanford, California, USA
[5]Department of Orthopedic Surgery and Department of Bioengineering, Stanford University, Stanford, California, USA

Correspondence to
Dr William H Robinson, Division of Immunology and Rheumatology, Stanford University School of Medicine, 269 Campus Dr, Stanford, CA 94305, USA; w.robinson@stanford.edu

HR and CML are co-first authors and contributed equally to this work.

Received 27 August 2016
Revised 21 October 2016
Accepted 19 November 2016
Published Online First 13 December 2016

To cite: Raghu H, Lepus CM, Wang Q, *et al*. *Ann Rheum Dis* 2017;**76**:914–922.

ABSTRACT

Objectives While various monocyte chemokine systems are increased in expression in osteoarthritis (OA), the hierarchy of chemokines and chemokine receptors in mediating monocyte/macrophage recruitment to the OA joint remains poorly defined. Here, we investigated the relative contributions of the CCL2/CCR2 versus CCL5/CCR5 chemokine axes in OA pathogenesis.

Methods *Ccl2*-, *Ccr2*-, *Ccl5*- and *Ccr5*-deficient and control mice were subjected to destabilisation of medial meniscus surgery to induce OA. The pharmacological utility of blocking CCL2/CCR2 signalling in mouse OA was investigated using bindarit, a CCL2 synthesis inhibitor, and RS-504393, a CCR2 antagonist. Levels of monocyte chemoattractants in synovial tissues and fluids from patients with joint injuries without OA and those with established OA were investigated using a combination of microarray analyses, multiplexed cytokine assays and immunostains.

Results Mice lacking CCL2 or CCR2, but not CCL5 or CCR5, were protected against OA with a concomitant reduction in local monocyte/macrophage numbers in their joints. In synovial fluids from patients with OA, levels of CCR2 ligands (CCL2, CCL7 and CCL8) but not CCR5 ligands (CCL3, CCL4 and CCL5) were elevated. We found that CCR2+ cells are abundant in human OA synovium and that CCR2+ macrophages line, invade and are associated with the erosion of OA cartilage. Further, blockade of CCL2/CCR2 signalling markedly attenuated macrophage accumulation, synovitis and cartilage damage in mouse OA.

Conclusions Our findings demonstrate that monocytes recruited via CCL2/CCR2, rather than by CCL5/CCR5, propagate inflammation and tissue damage in OA. Selective targeting of the CCL2/CCR2 system represents a promising therapeutic approach for OA.

INTRODUCTION

Osteoarthritis (OA) is the most common form of arthritis for which there are currently no disease-modifying therapies available.[1] A growing body of evidence suggests that chronic, low-grade inflammation involving both the innate and adaptive immune systems is critical to the pathogenesis of OA,[2–4] yet the precise underlying cellular and molecular mechanisms remain poorly understood. It is well documented that monocytes/macrophages infiltrate OA synovial tissues, and monocyte/macrophage-derived inflammatory cytokines and chemokines are elevated in OA synovial tissues and fluids.[5–7] Moreover, radiographic OA severity and joint symptoms have been shown to correlate with the number of activated macrophages in the synovial tissue of individuals with knee OA.[8] However, the principal mechanisms of monocyte recruitment to the joint, namely differential contributions of specific chemokine–chemokine receptor axes, remain unclear.

While monocyte recruitment in the context of acute infectious inflammation has been extensively studied, relatively less is known about monocyte recruitment in chronic inflammatory diseases (eg, atherosclerosis).[9 10] The CC chemokine receptors CCR2 and CCR5 as well as their cognate ligands (eg, CCL2, CCL7 and CCL8 for CCR2, and CCL5, CCL3 or CCL4 for CCR5) have been shown to modulate monocyte/macrophage recruitment in multiple inflammatory diseases.[9] Indeed, several chemokines that mediate monocyte/macrophage and their receptors are detected in OA synovial tissues and synovial fluids.[11 12] Additional studies have reported that synovium from individuals with joint injuries such as meniscal tears, a major risk factor for OA development, have elevated chemokine expression suggesting that these chemokines might instigate inflammatory responses in such individuals.[13 14] Furthermore, a recent study reported that synovial fluid levels of CCL2 (or MCP-1) correlated positively with pain and physical disability in patients with OA.[15] In line with this, it has been shown that deficiency in CCR2, the main receptor for CCL2, reduces OA-related pain mouse OA.[16] Nevertheless, the functional involvement of CCL2 in the OA disease process and the relative contribution of the CCL2/CCR2 versus the CCL5/CCR5 chemokine systems to monocyte recruitment in OA remains unknown. Understanding the processes, driving monocyte recruitment could aid in the development of novel targeted therapies to selectively inhibit pathological responses in OA.

METHODS

Study approval

We studied human samples under protocols approved by the Stanford Institutional Review

Board (IRB) and the University of Padova IRB, and with the subjects' informed consent.

Animals

C57BL/6J, *Ccr2*$^{-/-}$ (B6.129S4-*Ccr2*$^{tm1Ifc/J}$), *Ccl2*$^{-/-}$ (B6.129S4-*Ccl2*$^{tm1Rol/J}$), *Ccr5*$^{-/-}$ (B6.129P2-*Ccr5*$^{tm1Kuz/J}$) and *Ccl5*$^{-/-}$ (B6.129P2-*Ccl5*$^{tm1Hso/J}$) mice were purchased from The Jackson Laboratory. Destabilisation of the medial meniscus (DMM) was performed as described previously.[17–19] All animal studies were performed under protocols approved by the Stanford Committee of Animal Research and in accordance with National Institutes of Health guidelines.

Statistics

Data were analysed using two-tailed Student's t-test or Mann-Whitney U test for parametric and non-parametric data, respectively. $p<0.05$ was considered statistically significant.

Detailed methods are described in online supplementary materials.

RESULTS

CCL2 deficiency protects against mouse OA and is associated with reduced synovial macrophages and inflammatory mediators

Recently, it was reported that levels of CCL2, a major monocyte chemoattractant in serum and synovial fluid obtained from individuals with OA correlated positively with pain and physical disability.[15] Further, microarray analyses of genes expressed in knees of mice with surgery-induced OA showed that *Ccl2* is upregulated as early as 6 hours and 7 days after surgery when there is no apparent histological changes) relative to sham-operated controls (see online supplementary figure S1A).[18 20] To evaluate whether increased *Ccl2* gene expression was sustained through the progression of OA, we analysed synovial *Ccl2* expression at 12 weeks after DMM, when significant cartilage degradation is apparent.[18 19 21] We found that synovial *Ccl2* mRNA was significantly upregulated in DMM-operated compared with sham-operated mice (see online supplementary figure S1B). Notably, mining analyses of microarray data from knees of STR/ort mice, which spontaneously develop progressive OA similar to human OA starting at roughly 12 weeks of age,[22] also revealed significant upregulation of *Ccl2* relative to similarly aged OA-resistant CBA mice (see online supplementary figure S1C). Together, these data suggest that CCL2 might be involved in mechanical trauma-induced OA and in ageing-related spontaneous OA.

To directly demonstrate a pathogenic role for CCL2 in OA, we surgically induced OA via DMM in CCL2-deficient (*Ccl2*−/−) and age-matched wild-type (WT) mice. Histological analyses of knee joints 20 weeks after DMM revealed that *Ccl2*−/− mice exhibit less severe cartilage damage (figure 1A, B), osteophyte formation (figure 1A and online supplementary figure S2B) and synovitis (see online supplementary figures S2A, B) compared with WT controls. Consistent with CCL2's function as a potent monocyte chemoattractant, we found that macrophage numbers were significantly reduced in knee joints of *Ccl2*−/− mice compared with WT mice (figure 1C, D). Furthermore, analysis of synovial mRNA revealed significant reductions in inflammatory and degradative enzyme expression in *Ccl2*−/− compared with WT mice (figure 1E), suggesting that CCL2 deficiency downregulates local inflammatory responses in experimental mouse OA.

Reduced mouse OA and local macrophage numbers in CCR2-deficient mice

Because CCR2 is the main receptor for CCL2 and CCR2-induced monocyte recruitment has been shown to mediate pain in the DMM model,[16] we tested whether CCR2-mediated monocyte recruitment contributes to the development of mouse OA. Whereas WT mice developed severe cartilage damage after DMM surgery, *Ccr2*−/− mice exhibited significantly less cartilage damage (figure 2A, B), synovitis (see online supplementary figures S3A, B) and osteophyte formation (see online supplementary figure S3B). The role of CCR2 in

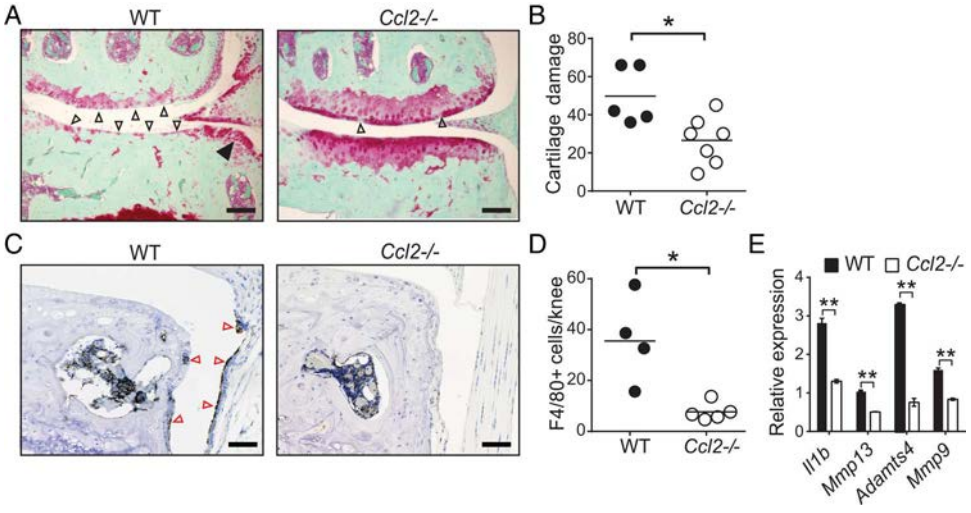


Figure 1 CCL2 deficiency protects against development of osteoarthritis in mice, and is associated with reduced synovial macrophages and inflammatory mediators. (A) Representative safranin-o stained knee joint sections showing extensive cartilage damage (open arrows) and osteophytes (filled arrows) and (B) quantification of cartilage damage in wild-type (WT, n=5) but not *Ccl2*−/− (n=7) mice 20 weeks after destabilisation of the medial meniscus (DMM) surgery by a blinded investigator. (C, D) Representative immunostains and quantification of F4/80+ (brown) macrophages in the synovium (red arrows) in WT and *Ccl2*−/− mice 16 weeks after DMM surgery. Symbols represent individual mice and bars denote the mean in (B) and (D). (E) Quantitative PCR (qPCR) analyses of proinflammatory gene expression in synovium of WT (n=3) and *Ccl2*−/− (n=3) mice 10 weeks after DMM surgery. qPCR data are mean±SEM. Scale bars 200 μm. *p<0.05, **p<0.01 by the Mann-Whitney U test for (B), (D) and by Student's t-test for (E). The presented data are representative of two independent experiments with similar results.

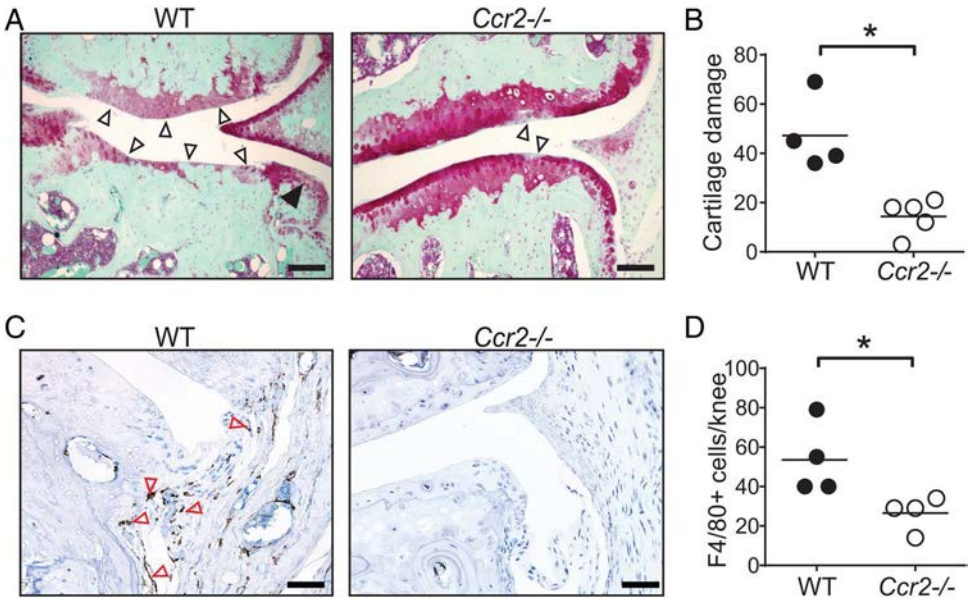


Figure 2 CCR2 deficiency attenuates osteoarthritis in mice. (A) Representative safranin-o stained knee joint sections showing severe cartilage damage (open arrows) and osteophytes (filled arrows) and (B) quantification of cartilage damage in wild-type (WT, n=4) but not *Ccr2*−/− (n=5) mice 20 weeks after destabilisation of the medial meniscus surgery by a blinded investigator. (C) Representative immunohistochemistry and (D) quantification of F4/80+ (brown) macrophages lining the synovium (red arrows) in WT and *Ccr2*−/− mice. Symbols denote individual mice, and bars denote the mean of indicated groups in (B) and (D). Scale bars denote 200 μm. *p<0.05 by the Mann-Whitney U test for (B) and by Student's t-test for (D). The presented data are representative of two independent experiments with similar results.

macrophage recruitment was confirmed by immunohistochemical analysis: whereas the knee joints of WT mice contained substantial numbers of macrophages after DMM, those of *Ccr2*−/− mice were largely devoid of such cells (figure 2C, D). Together, these findings demonstrate that both CCL2 and CCR2 promote inflammation and tissue damage in OA via mechanism(s) linked to monocyte/macrophage recruitment.

CCR5 or CCL5 deficiencies do not attenuate experimental mouse OA development or diminish macrophage burden in the OA knee joints

It has been well documented that while CCL2 is a potent monocyte chemoattractant, other chemokine–chemokine receptor systems such as CCL5/CCR5 are also involved in monocyte recruitment during inflammation.[9 10] To further investigate the relative contribution of the CCR5/CCL5 chemokine axis in mouse OA, we surgically induced OA in CCR5 or CCL5-deficient mice and found that 20 weeks after DMM surgery, cartilage damage was indistinguishable between CCR5-deficient (*Ccr5*−/−) and WT mice (figure 3A, B). Remarkably, we found no differences in macrophage numbers between WT and *Ccr5*−/− mice (figure 3C, D). Similar to mice deficient in CCR5, we found that mice deficient in CCL5, a major CCR5 ligand, also develop severe OA similar to that observed in WT controls (figure 3E, F) and have no apparent reduction in macrophage numbers in their knee joints (figure 3G, H). Together, our data indicate that neither CCL5 nor CCR5 is required for the development of OA in mice.

Increased monocyte chemoattractant proteins in human OA synovial tissue and fluid

In human OA, levels of various chemokines and chemokine receptors are abnormally high in the synovial fluid and synovial tissue, but little is known about the relative contribution of these molecules to OA pathology.[11 23] Unsupervised cluster analyses of publicly available gene expression dataset GSE32317 revealed that multiple chemokines involved in monocyte recruitment or macrophage accumulation are overexpressed in synovial membranes derived from individuals with early or end-stage OA as compared with membranes from 'healthy' joints (figure 4A). To validate this, we performed analyses of another independent dataset (GSE46750) and demonstrated that the same putative pathogenic monocyte chemokine encoding genes (eg, *CCL2*, *CCL5*, *CCL19* and *CXCL8*) were also statistically upregulated in inflamed OA synovium as compared with non-inflamed OA synovium (see online supplementary table S1).

Chemokines that signal via the CCR2 receptor are elevated in the synovial fluids of individuals with OA

Next, we analysed protein levels of various chemokines in synovial fluids of those with established OA, with rheumatoid arthritis (RA) or obtained from individuals with post-traumatic joint injuries that are at increased risk for developing OA[24] (labelled 'PT non-OA'). Consistent with the concept that OA is associated with low-grade inflammation,[25 26] we found that chemokine levels were elevated in OA synovial fluids compared with PT non-OA samples but generally lower than the levels in RA synovial fluids (figure 4B, C and online supplementary figures S7A, B). Importantly, the levels of multiple CCR2 ligands such as CCL2 (figure 4B), CCL7 (see online supplementary figure S4B) and CCL8 (see online supplementary figure S4C) were significantly elevated in OA synovial fluids. By contrast, OA levels of the major CCR5 ligands CCL5 (figure 4C) and CCL3 (see online supplementary figure S4D) were either undetectable or very low in these very same synovial fluids.

CCL2 but not CCL5 is secreted by human OA synovial fibroblasts on stimulation with cartilage debris

CCL2 is predominantly produced by immune cells in humans and mice. In the context of OA, previous studies demonstrated

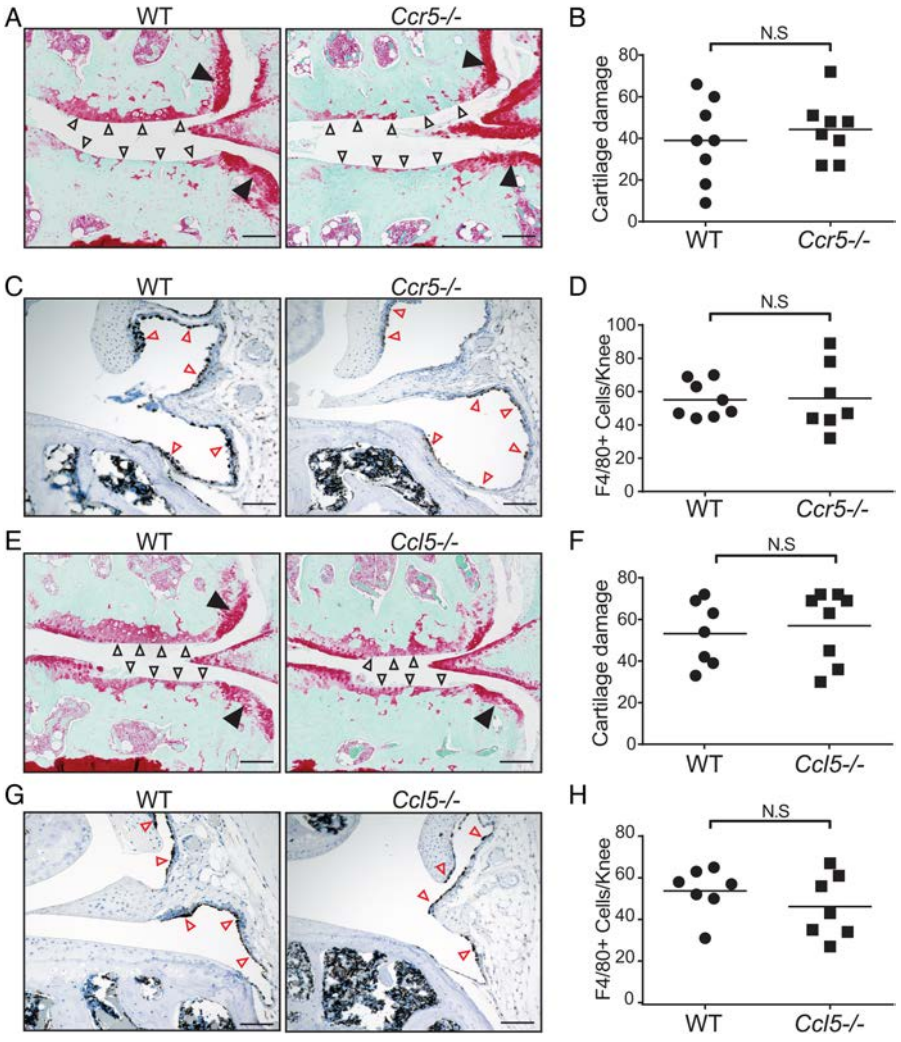


Figure 3 CCR5 or CCL5 deficiency does not modulate the severity of experimental osteoarthritis following destabilisation of the medial meniscus (DMM). (A) Representative safranin-o stained knee joint sections showing severe cartilage damage (open arrows) and osteophytes (filled arrows) in both wild-type (WT) and *Ccr5*−/− mice 20 weeks after DMM surgery. (B) Quantification of cartilage damage in WT (n=8) and *Ccr5*−/− (n=8) mice by a blinded investigator. (C) Representative immunohistochemistry showing F4/80+ (brown) macrophages lining the synovium (red arrows) in WT and *Ccr5*−/− mice. (D) Quantification of F4/80+ macrophages in WT and *Ccr5*−/− mice. (E) Representative safranin-o stained knee joint sections showing marked cartilage damage (open arrows) and osteophytes (filled arrows) in WT and *Ccl5*−/− mice 20 weeks after DMM surgery. (F) Quantification of cartilage damage in WT (n=7) and *Ccl5*−/− (n=8) mice by a blinded investigator. (G) Representative immunohistochemistry showing F4/80+ (brown) macrophages lining the synovium (red arrows) in WT and *Ccl5*−/− mice. (H) Quantification of F4/80+ macrophages in WT and *Ccl5*−/− mice. Symbols denote individual mice, and bars denote the mean. Scale bars denote 200 μm. *p<0.05 by the Mann-Whitney U test in (B) and (F) and Student's t-test in (D) and (H).

that chondrocytes upregulate CCL2 expression following inflammatory stimuli.[17 27 28] Consistent with this, we found that chondrocytes within OA cartilage tissues abundantly express CCL2 (see online supplementary figure S4E). We next tested whether synovial fibroblasts might also produce CCL2 in response to tissue injury by stimulating OA synovial fibroblasts with OA cartilage-derived debris, or the alarmin S100A8. First, we observed that unstimulated OA synovial fibroblasts abundantly secreted CCL2 (figure 4D), but not CCL5 (figure 4E). Next, we found that while synovial fibroblasts stimulated with cartilage debris or S100A8 secreted CCL2 (figure 4D), only stimulation with S100A8 induced robust CCL5 secretion (figure 4E). Finally, we found that CCL2 production by both unperturbed and stimulated fibroblasts could be effectively limited using bindarit, a specific CCL2 synthesis inhibitor (figure 4D). Consistent with previous reports showing that bindarit is highly specific to the monocyte chemoattractant protein (MCP) subfamily which includes CCL2,[29] bindarit did not mitigate CCL5 secretion from synovial fibroblasts stimulated with S100A8 (figure 4E).

CCR2+ macrophages are abundant in OA synovium and in close physical association with articular cartilage tissues

To further analyse whether CCL2 signalling is upregulated in OA synovial tissue, we performed analyses of the GSE32317 dataset and found that gene expression of CCR2 was also upregulated in OA synovium relative to healthy synovium (figure 5A). Supporting this observation, immunofluorescent analysis showed that the number of CCR2-expressing cells (CCR2+) was also significantly higher in OA as compared with PT non-OA samples (figure 5B, C), suggesting an important role for CCR2-expressing cells in human OA pathogenesis. It is well described that CCR2-expressing inflammatory monocytes recruited to tissues differentiate into proinflammatory macrophages and drive local inflammation and tissue damage.[30–32] Thus, we hypothesised that CCR2+ macrophages could be directly involved in OA tissue damage. In agreement with this, we found many CCR2-expressing cells in close proximity to the surface of articular cartilage (figure 5D, green). Further, we found that these CCR2+ cells are haematopoietic in origin as they also expressed CD45 (data not shown) and are indeed macrophages as evidenced by CD68 staining (figure 5D, red). Importantly, as seen in online supplementary figure S5, this inflammatory layer of cells was found adjacent to articular cartilage surfaces. However, further analyses needed to define their precise role in cartilage degradation.

Pharmacological intervention at the level of CCL2 synthesis or its binding to CCR2 attenuates mouse OA development and severity

Based on our findings that genetic deficiency of CCL2 attenuates local inflammation and tissue damage in experimental OA and that the CCL2/CCR2 axis is upregulated in human OA, we evaluated the utility of blocking CCL2 synthesis in mouse OA using bindarit, a molecule previously shown to inhibit CCL2 induction in vivo.[33 34] While vehicle-treated mice developed

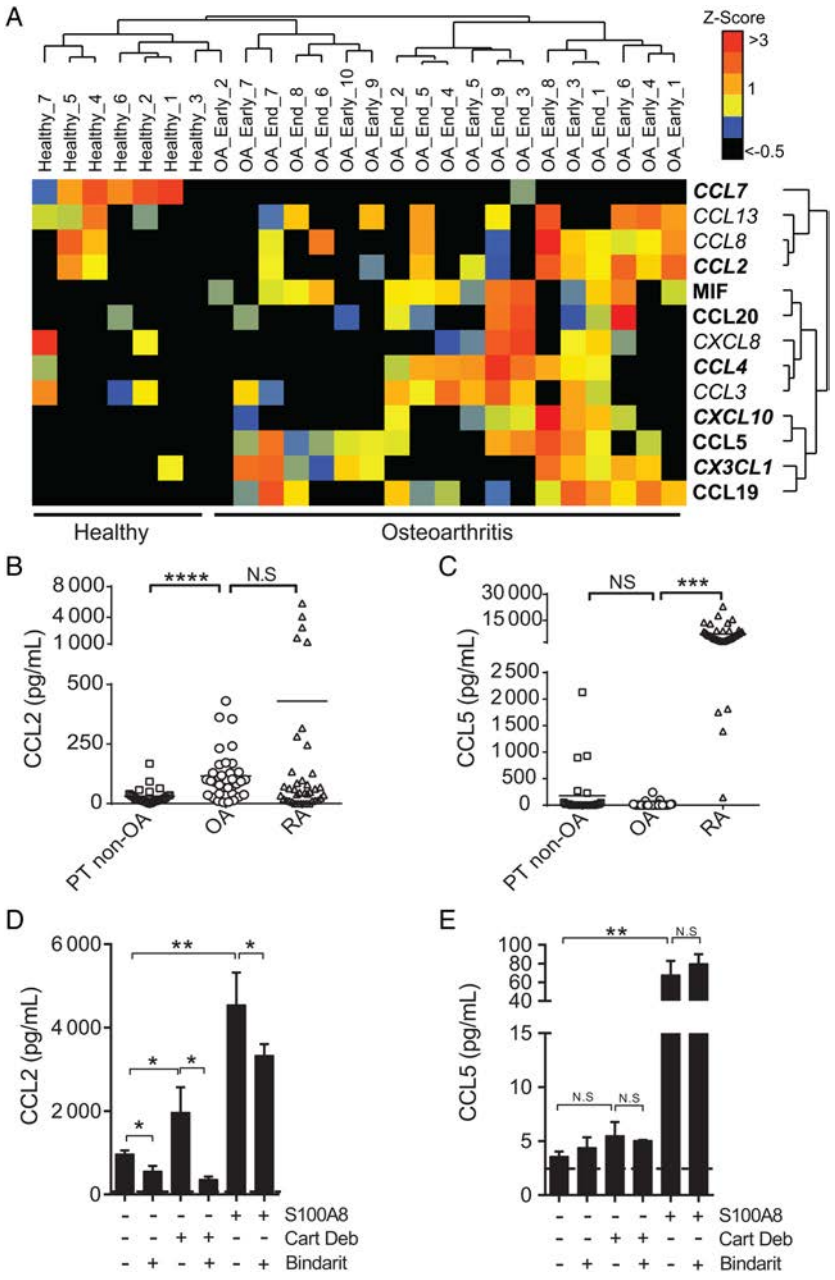


Figure 4 Enhanced monocyte chemoattractant mRNA expression in human osteoarthritis (OA) synovial tissue and increased CCL2, but not CCL5, levels in human OA synovial fluids and synovial fibroblasts. (A) Unsupervised cluster analyses of gene expression of various chemokines involved in monocyte chemoattraction in microarray datasets from synovial membranes of individuals with prior traumatic joint injury but no radiographic OA (healthy) and from individuals with early-stage or end-stage OA (downloaded from gene expression omnibus (GEO), accession code GSE32317). Scale represents Z scores. Bold font indicates significant difference between 'healthy' and 'osteoarthritis' groups by significance analysis of microarrays (SAM) analyses with q-value (false discovery rate) cut-off set at 0.05. Quantification of (B) CCL2 and (C) CCL5 in synovial fluids of individuals with posttraumatic joint injuries (PT non-OA, n=37), OA (n=35) or rheumatoid arthritis (n=26). Analyses of (D) CCL2 and (E) CCL5 levels in supernatants of human OA-derived primary synovial fibroblasts stimulated with 20 mg/mL OA cartilage debris (Cart Deb), 1 μg/mL S100A8 (positive control) or media alone in the presence or absence of 300 μM bindarit for 24 hours. In vitro stimulation assays were performed in triplicate and are representative of data obtained from four independent synovial fibroblast cultures derived from four different individuals with OA undergoing total knee arthroplasty. *p<0.05, **p<0.01, ****p<0.0001 Student's t-test, NS, not significant.

severe OA, mice treated with bindarit for 12 weeks exhibit markedly diminished cartilage damage (figure 6A, B), synovitis and osteophyte formation (see online supplementary figure S6A, B). In addition, bindarit treatment significantly reduced macrophage accumulation in mouse knee joints (figure 6C, D).

To further confirm a pathogenic role for CCL2/CCR2 signalling in the pathogenesis of mouse OA, we tested whether RS-504393, a CCR2 antagonist that potently inhibits CCL2 binding to CCR2 but not to other CCL2 receptors,[35] could attenuate OA disease progression and/or severity. We found that RS-504393 significantly diminished cartilage damage (figure 6E, F), synovitis and osteophyte formation (see online supplementary figures S7A,B). Indeed, the number of F4/80-positive macrophages was significantly lower in RS-504393-treated mice as compared with vehicle-treated mice (figure 6G, H). Thus, pharmacological blockade of the CCL2/CCR2 chemokine system effectively diminishes mouse OA in part by reducing synovial macrophage accumulation.

DISCUSSION

The objective of this study was to examine the differential involvement of chemokine–chemokine receptor systems in monocyte recruitment in OA using experimental mouse model systems and human OA tissues. Here, we report that the CCL2/CCR2 signalling axis preferentially mediates monocyte trafficking and promotes inflammation and tissue damage in OA. Importantly, we show that as in mouse OA, in human OA CCR2 ligands such as CCL2 and CCL8, but not CCR5 ligands such as CCL5, are preferentially elevated. Finally, we demonstrate for the first time that pharmacological inhibition of CCL2 synthesis or its binding to CCR2 protects against development of mouse OA in part by attenuating macrophage accumulation in the synovial joints.

In several mouse models of chronic inflammation (eg, chronic kidney disease, RA, asthma, etc), deficiencies in CCL2 or CCR2 protect against inflammation and tissue damage.[36] Previous reports also suggest that specific chemokine–chemokine receptor

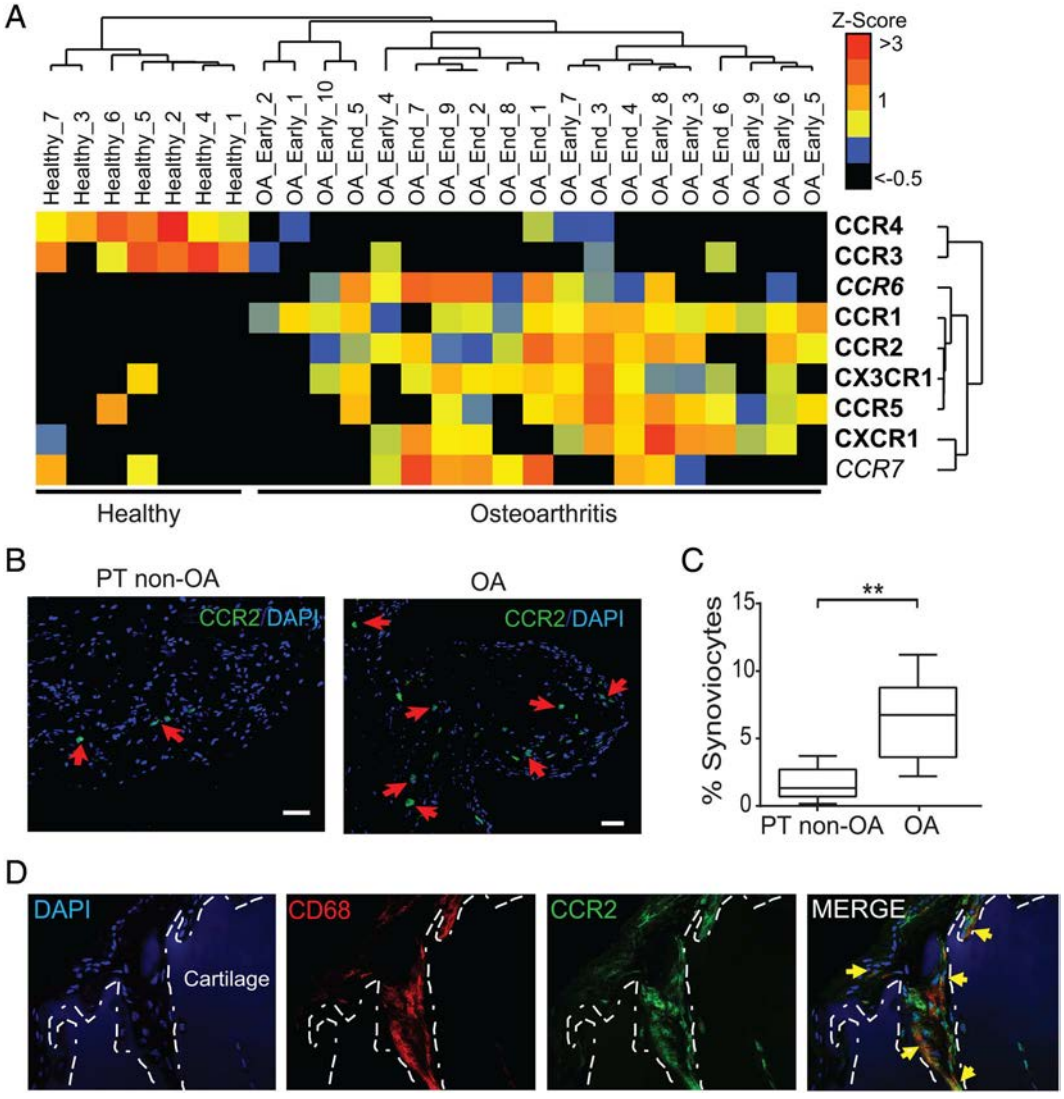


Figure 5 Increased mRNA expression of chemokine receptors and increased CCR2+ macrophages in synovium and cartilage tissues obtained from humans with osteoarthritis (OA). (A) Unsupervised cluster analyses of gene expression of various chemokine receptors involved in monocyte chemoattraction in microarray datasets from synovial membranes of individuals with prior traumatic joint injury but no radiographic OA (healthy) and from individuals with early-stage or end-stage OA (downloaded from GEO, accession code GSE32317). Scale represents Z scores. Bold font indicates significant difference between 'healthy' and 'osteoarthritis' groups by SAM analyses with q-value (false discovery rate) cut-off set at 0.05. (B) Representative immunofluorescent staining of CCR2 (red arrows) in sections of PT non-OA or OA synovial tissues. (C) Quantification of CCR2+ cells in synovial tissue as a percentage of total DAPI-stained synoviocytes per random low power field (LPF). Ten LPFs were quantitated per synovial tissue section (n=5 individual samples per group). Error bars indicate maximum and minimum values, and line denotes their mean. (D) Representative immunostaining of CD68 (red) and CCR2 (green) adjacent to the lesional articular cartilage samples from patients with OA (n=6) showing CD68+CCR2+ double-positive macrophages (yellow arrows) invading the cartilage tissue. Dotted white lines demarcate the cartilage tissue lined by invading macrophages. Scale bar denotes 200 μm. **p<0.01 Student's t-test.

systems may differentially contribute to monocyte recruitment in a time and context-dependent manner. For instance, in a model of high fat diet-induced atherosclerosis, CCR1 and CCR5 but not CCR2 or CX3CR1 were found to be crucial for monocyte recruitment.[31] In a setting of DSS-induced colitis in mice, both CCR2 and CCR5 were found to be important for driving inflammatory responses.[37] Here, we report that in OA there exists an apparent hierarchy in chemokine-driven mechanisms in that CCL2/CCR2-driven but not CCL5/CCR5-driven monocyte recruitment promotes mouse and human OA pathogenesis. CCR2 has additional ligands in humans and mice (eg, CCL7, CCL8, CCL13 and CCL14), but studies in mice lacking different combinations of these ligands suggest that CCL2 and CCL7 are critically involved in monocyte mobilisation from the bone marrow.[38] Nevertheless, it remains to be experimentally determined whether these additional CCR2 ligands, as well as other CCL2 receptors such as CCR1 and CCR4, modulate the pathogenesis of OA. Further, while monocyte/macrophage accumulation within the inflamed synovial joints is thought to be fundamentally critical for OA pathogenesis, other tightly controlled processes such as survival, activation and polarisation could also be crucial in modulating OA pathogenesis.

The chemokine CCL5 and its receptor CCR5 also possess strong monocyte chemoattractive properties. Here, we show that neither CCR5 nor CCL5 deficiency confers protection against the development or severity of OA in mice. This finding is in direct contrast to prior findings by Takebe *et al*[39] showing reduced cartilage damage in *Ccr5*−/− mice. These discrepancies

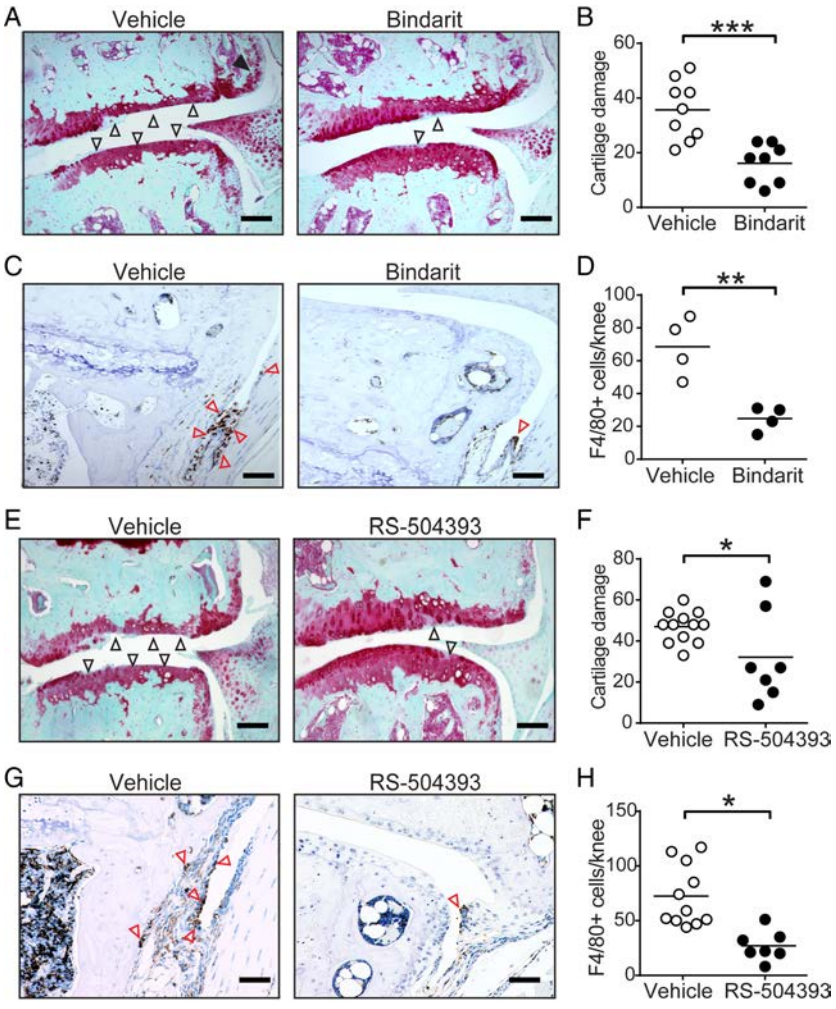


Figure 6 Pharmacological blockage of CCL2 synthesis or binding to CCR2 protects against osteoarthritis development in mice. (A) Representative safranin-o stained knee joint sections and (B) quantification of cartilage damage from mice receiving vehicle (n=10) or 50 mg/kg/day bindarit (n=8) by oral gavage for 12 weeks after destabilisation of the medial meniscus (DMM) surgery. Open and filled arrows indicate cartilage damage and osteophytes, respectively. (C) Representative immunostains and (D) quantification of F4/80+ (brown) macrophages in the synovium (red arrows) of vehicle-treated but not bindarit-treated mice. (E) Representative safranin-o stained knee joint sections and (F) quantification of cartilage damage from mice receiving vehicle (n=12) or 4 mg/kg/day RS-504393 (n=7) by oral gavage for 12 weeks after DMM surgery. Open and filled arrows indicate cartilage damage and osteophytes, respectively. (G) Representative immunostains and (H) quantification of F4/80+ (brown) macrophages in the synovium (red arrows) of vehicle-treated but not RS-504393-treated mice. Symbols denote individual mice and bars denote the mean of indicated groups in (B) and (F). Scale bars denote 200 μm. *p<0.05 by the Mann-Whitney U test for (B) and (F) and by Student's t-test for (D) and (H).

can potentially be attributed to differences in experimental design, including (1) age at OA induction (20-week-old mice used in our studies vs 10-week-old mice in studies by Takebe *et al*,[39] as older mice develop more severe OA),[40 41] (2) duration of OA development (DMM-associated pathologies were assessed 20 weeks after surgery, whereas Takebe *et al*[39] characterised changes at 8 and 12 weeks after DMM) and (3) possible sex-related differences in disease severity (male mice, used in our study, suffer more OA pathology compared with female mice.[41 42] It is unclear which sex was used by Takebe *et al*).[39]

Identifying the source of inflammatory chemokines such as CCL2 is fundamentally important to understanding disease biology and developing therapeutic interventions. However, the cell types responsible for CCL2 secretion in OA remain unclear. While it is known that CCL2 is predominantly made by immune cells in humans and mice, other non-immune cell types have also been shown to express CCL2. For instance, previous studies have demonstrated that cultured murine and human chondrocytes upregulate CCL2 expression following inflammatory stimuli.[27 28] Furthermore, it is well known that ageing, a major risk factor for OA development, is associated with systemic low-grade inflammation including elevated local and circulating cytokines.[43] In this regard, senescent cells such as ageing articular chondrocytes release cytokines including CCL2 that could contribute to inflammatory mechanisms that promote OA as shown in the studies presented here.[44] Here, we report that CCL2 is abundantly secreted by human OA synovial fibroblasts on stimulation with cartilage debris. Based on our observations, we propose a model where joint-resident cells such as chondrocytes and/or synovial fibroblasts are the initial source of CCL2 following injury, thereby recruiting CCR2-expressing inflammatory monocytes that propagate OA pathogenesis via production of inflammatory cytokines and tissue degradative enzymes.

We show that CCL2 secretion can be abrogated using bindarit, a synthetic molecule previously shown to potently and selectively inhibit expression of the monocyte chemoattractants CCL2, CCL8 and CCL7 as well as attenuate inflammation in various mouse models of inflammatory disease including experimental autoimmune encephalomyelitis (EAE) and 2,4,6-trinitrobenzenesulfonic acid (TNBS)-induced colitis.[29 34] We provide proof of concept that pharmacological suppression of CCL2 expression using bindarit is effective in reducing inflammation, cartilage damage and monocyte/macrophage accumulation in mouse OA. We further demonstrate that pharmacological blockage of CCR2 signalling using RS-504393, a CCR2 antagonist that has been shown to specifically inhibit CCL2/CCR2 interactions,[35] significantly attenuates OA disease progression and/or severity in mice through a mechanism linked to synovial macrophage accumulation. Notably, RS-504393 has been previously demonstrated to reduce pain in this mouse model in part by inhibiting macrophage recruitment to the dorsal root ganglion. Given the critical role of the CCL2/CCR2 system in recruiting inflammatory monocytes, long-term systemic blockade of CCL2/CCR2 could potentially increase susceptibility to infections, impair wound healing and tissue repair.

Thus, local inhibition of CCL2 or CCR2 might be a more effective and safe strategy for the treatment of OA.

Indeed, the prominent role of the CCL2/CCR2 chemokine system in inflammation has rendered it an attractive target for therapeutic intervention in multiple diseases including RA. However, pharmacological inhibition of CCL2 binding to CCR2 has failed in clinical trials to date, most conspicuously in RA.[45–47] The reasons for failure are likely multifold, have been reviewed elsewhere and can be in part attributed to off-target effects on CCR5.[48 49] It is noteworthy that OA immunopathogenesis is distinct from that of RA, including the use of distinct chemokine–chemokine receptor systems for the recruitment of monocytes/macrophages to the 'low-grade' inflamed OA joint. Based on our results, we propose that selective targeting of the CCL2/CCR2 system, either alone or in combination with other therapies, has the potential to provide therapeutic benefit in OA.

Acknowledgements The authors thank Dr T Lindstrom for critically reviewing and editing the manuscript and Dr D Rajamani for contributing to microarray data analysis.

Contributors HR and CML performed key studies. HR, QW and HHW conducted the studies of osteoarthritis in mouse models. CML and HR performed immunofluorescent and H&E staining of synovium provided by NJG, SBG and CRC. HR analysed these images. HR and NL conducted ELISA analyses, and NL performed Luminex cytokine profiling using synovial fluids provided by JBS, CRC, LP and FO. HR cultured and performed the in vitro stimulation assays on OA synovial fibroblasts. HR and NL analysed gene expression datasets downloaded from NCBI. CRC and JBS provided key scientific input. HR and WHR wrote and edited the manuscript. All authors reviewed the data and approved the manuscript.

Funding These studies were supported by VA RR&D Merit Review Awards I01BX002345, I01RX000934 and I01RX000588 to WHR.

Competing interests None.

Ethics approval Stanford Institutional Review Board, University of Padova IRB.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Felson DT. Clinical practice. Osteoarthritis of the knee. *N Engl J Med* 2006;354:841–8.
2 Liu-Bryan R, Terkeltaub R. Emerging regulators of the inflammatory process in osteoarthritis. *Nat Rev Rheumatol* 2015;11:35–44.
3 Klareskog L, Johnell O, Hulth A, *et al*. Reactivity of monoclonal anti-type II collagen antibodies with cartilage and synovial tissue in rheumatoid arthritis and osteoarthritis. *Arthritis Rheum* 1986;29:730–8.
4 Goldring MB, Otero M. Inflammation in osteoarthritis. *Curr Opin Rheumatol* 2011;23:471–8.
5 Kapoor M, Martel-Pelletier J, Lajeunesse D, *et al*. Role of proinflammatory cytokines in the pathophysiology of osteoarthritis. *Nat Rev Rheumatol* 2011;7:33–42.
6 Rahmati M, Mobasheri A, and Mozafari M. Inflammatory mediators in osteoarthritis: a critical review of the state-of-the-art, current prospects, and future challenges. *Bone* 2016;85:81–90.
7 Berenbaum F. Osteoarthritis as an inflammatory disease (osteoarthritis is not osteoarthrosis!). *Osteoarthr Cartil* 2013;21:16–21.
8 Kraus VB, McDaniel G, Huebner JL, *et al*. Direct in vivo evidence of activated macrophages in human osteoarthritis. *Osteoarthr Cartil* 2016;24:1613–21.
9 Griffith JW, Sokol CL, and Luster AD. Chemokines and chemokine receptors: positioning cells for host defense and immunity. *Annu Rev Immunol* 2014;32:659–702.
10 Shi C, Pamer EG. Monocyte recruitment during infection and inflammation. *Nat Rev Immunol* 2011;11:762–74.
11 Endres M, Andreas K, Kalwitz G, *et al*. Chemokine profile of synovial fluid from normal, osteoarthritis and rheumatoid arthritis patients: CCL25, CXCL10 and XCL1 recruit human subchondral mesenchymal progenitor cells. *Osteoarthr Cartil* 2010;18:1458–66.
12 Haringman JJ, Smeets TJ, Reinders-Blankert P, *et al*. Chemokine and chemokine receptor expression in paired peripheral blood mononuclear cells and synovial tissue of patients with rheumatoid arthritis, osteoarthritis, and reactive arthritis. *Ann Rheum Dis* 2006;65:294–300.
13 Scanzello CR, McKeon B, Swaim BH, *et al*. Synovial inflammation in patients undergoing arthroscopic meniscectomy: molecular characterization and relationship to symptoms. *Arthritis Rheum* 2011;63:391–400.
14 Nair A, Gan J, Bush-Joseph C, *et al*. Synovial chemokine expression and relationship with knee symptoms in patients with meniscal tears. *Osteoarthr Cartil* 2015;23:1158–64.
15 Li L, and Jiang BE. Serum and synovial fluid chemokine ligand 2/monocyte chemoattractant protein 1 concentrations correlates with symptomatic severity in patients with knee osteoarthritis. *Ann Clin Biochem* 2015;52(Pt 2):276–82.
16 Miller RE, Tran PB, Das R, *et al*. CCR2 chemokine receptor signaling mediates pain in experimental osteoarthritis. *Proc Natl Acad Sci USA* 2012;109:20602–7.
17 Wang Q, Rozelle AL, Lepus CM, *et al*. Identification of a central role for complement in osteoarthritis. *Nat Med* 2011;17:1674–9.
18 Loeser RF, Olex AL, McNulty MA, *et al*. Disease progression and phasic changes in gene expression in a mouse model of osteoarthritis. *PLoS ONE* 2013;8:e54633.
19 Glasson SS, Askew R, Sheppard B, *et al*. Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis. *Nature* 2005;434:644–8.
20 Burleigh A, Chanalaris A, Gardiner MD, *et al*. Joint immobilization prevents murine osteoarthritis and reveals the highly mechanosensitive nature of protease expression in vivo. *Arthritis Rheum* 2012;64:2278–88.
21 Xu L, Servais J, Polur I, *et al*. Attenuation of osteoarthritis progression by reduction of discoidin domain receptor 2 in mice. *Arthritis Rheum* 2010;62:2736–44.
22 Kyostio-Moore S, Nambiar B, Hutto E, *et al*. STR/ort mice, a model for spontaneous osteoarthritis, exhibit elevated levels of both local and systemic inflammatory markers. *Comp Med* 2011;61:346–55.
23 Ritter SY, Subbaiah R, Bebek G, *et al*. Proteomic analysis of synovial fluid from the osteoarthritic knee: comparison with transcriptome analyses of joint tissues. *Arthritis Rheum* 2013;65:981–92.
24 Lotz MK, Kraus VB. New developments in osteoarthritis. Posttraumatic osteoarthritis: pathogenesis and pharmacological treatment options. *Arthritis Res Ther* 2010;12:211.
25 Kelso EB, Ferrell WR, Lockhart JC, *et al*. Expression and proinflammatory role of proteinase-activated receptor 2 in rheumatoid synovium: ex vivo studies using a novel proteinase-activated receptor 2 antagonist. *Arthritis Rheum* 2007;56:765–71.
26 Pessler F, Dai L, Diaz-Torne C, *et al*. The synovitis of "non-inflammatory" orthopaedic arthropathies: a quantitative histological and immunohistochemical analysis. *Ann Rheum Dis* 2008;67:1184–7.
27 Appleton CT, Usmani SE, Pest MA, *et al*. Reduction in disease progression by inhibition of transforming growth factor alpha-CCL2 signaling in experimental posttraumatic osteoarthritis. *Ann Rheum Dis* 2015;67:2691–701.
28 Xu YK, Ke Y, Wang B, and, *et al* The role of MCP-1-CCR2 ligand-receptor axis in chondrocyte degradation and disease progress in knee osteoarthritis. *Biol Res* 2015;48:64.
29 Mirolo M, Fabbri M, Sironi M, *et al*. Impact of the anti-inflammatory agent bindarit on the chemokinome: selective inhibition of the monocyte chemotactic proteins. *Eur Cytokine Netw* 2008;19:119–22.
30 Swirski FK, Libby P, Aikawa E, *et al*. Ly-6Chi monocytes dominate hypercholesterolemia-associated monocytosis and give rise to macrophages in atheromata. *J Clin Invest* 2007;117:195–205.
31 Tacke F, Alvarez D, Kaplan TJ, *et al*. Monocyte subsets differentially employ CCR2, CCR5, and CX3CR1 to accumulate within atherosclerotic plaques. *J Clin Invest* 2007;117:185–94.
32 Dal-Secco D, Wang J, Zeng Z, *et al*. A dynamic spectrum of monocytes arising from the in situ reprogramming of CCR2+ monocytes at a site of sterile injury. *J Exp Med* 2015;212:447–56.
33 Shen JZ, Morgan J, Tesch GH, *et al*. CCL2-dependent macrophage recruitment is critical for mineralocorticoid receptor-mediated cardiac fibrosis, inflammation, and blood pressure responses in Male mice. *Endocrinology* 2014;155:1057–66.
34 Ge S, Shrestha B, Paul D, *et al*. The CCL2 synthesis inhibitor bindarit targets cells of the neurovascular unit, and suppresses experimental autoimmune encephalomyelitis. *J Neuroinflammation* 2012;9:171.
35 Mirzadegan T, Diehl F, Ebi B, *et al*. Identification of the binding site for a novel class of CCR2b chemokine receptor antagonists: binding to a common chemokine receptor motif within the helical bundle. *J Biol Chem* 2000;275:25562–71.
36 Pérez de Lema G, Maier H, Franz TJ, *et al*. Chemokine receptor Ccr2 deficiency reduces renal disease and prolongs survival in MRL/lpr lupus-prone mice. *J Am Soc Nephrol* 2005;16:3592–601.
37 Andres PG, Beck PL, Mizoguchi E, *et al*. Mice with a selective deletion of the CC chemokine receptors 5 or 2 are protected from dextran sodium sulfate-mediated colitis: lack of CC chemokine receptor 5 expression results in a NK1.1+ lymphocyte-associated Th2-type immune response in the intestine. *J Immunol* 2000;164:6303–12.
38 Tsou CL, Peters W, Si Y, *et al*. Critical roles for CCR2 and MCP-3 in monocyte mobilization from bone marrow and recruitment to inflammatory sites. *J Clin Invest* 2007;117:902–9.
39 Takebe K, Rai MF, Schmidt EJ, *et al*. The chemokine receptor CCR5 plays a role in post-traumatic cartilage loss in mice, but does not affect synovium and bone. *Osteoarthr Cartil* 2015;23:454–61.

40 Loeser RF, Olex AL, McNulty MA, *et al*. Microarray analysis reveals age-related differences in gene expression during the development of osteoarthritis in mice. *Arthritis Rheum* 2012;64:705–17.
41 Fang H, Beier F. Mouse models of osteoarthritis: modelling risk factors and assessing outcomes. *Nat Rev Rheumatol* 2014;10:413–21.
42 Ma HL, Blanchet TJ, Peluso D, *et al*. Osteoarthritis severity is sex dependent in a surgical mouse model. *Osteoarthr Cartil* 2007;15:695–700.
43 Calçada D, Vianello D, Giampieri E, *et al*. The role of low-grade inflammation and metabolic flexibility in aging and nutritional modulation thereof: a systems biology approach. *Mech Ageing Dev* 2014;136–137:138–47.
44 Loeser RF, Collins JA, Diekman BO. Ageing and the pathogenesis of osteoarthritis. *Nat Rev Rheumatol* 2016;12:412–20.
45 Haringman JJ, Kraan MC, Smeets TJ, *et al*. Chemokine blockade and chronic inflammatory disease: proof of concept in patients with rheumatoid arthritis. *Ann Rheum Dis* 2003;62:715–21.
46 Haringman JJ, Gerlag DM, Smeets TJ, *et al*. A randomized controlled trial with an anti-CCL2 (anti-monocyte chemotactic protein 1) monoclonal antibody in patients with rheumatoid arthritis. *Arthritis Rheum* 2006;54:2387–92.
47 Vergunst CE, Gerlag DM, Lopatinskaya L, *et al*. Modulation of CCR2 in rheumatoid arthritis: a double-blind, randomized, placebo-controlled clinical trial. *Arthritis Rheum* 2008;58:1931–9.
48 Szekanecz Z, Koch AE. Successes and failures of chemokine-pathway targeting in rheumatoid arthritis. *Nat Rev Rheumatol* 2016;12:5–13.
49 Schall TJ, Proudfoot AE. Overcoming hurdles in developing successful drugs targeting chemokine receptors. *Nat Rev Immunol* 2011;11:355–63.

EXTENDED REPORT

Endothelial-to-mesenchymal transition contributes to endothelial dysfunction and dermal fibrosis in systemic sclerosis

Mirko Manetti,[1] Eloisa Romano,[2] Irene Rosa,[1,2] Serena Guiducci,[2] Silvia Bellando-Randone,[2] Amato De Paulis,[3] Lidia Ibba-Manneschi,[1] Marco Matucci-Cerinic[2]

Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210229).

[1]Department of Experimental and Clinical Medicine, Section of Anatomy and Histology, University of Florence, Florence, Italy
[2]Department of Experimental and Clinical Medicine, Section of Internal Medicine, Rheumatology Unit, Azienda Ospedaliero-Universitaria Careggi (AOUC), University of Florence, Florence, Italy
[3]Department of Translational Medical Sciences and Center for Basic and Clinical Immunology Research, University of Naples Federico II, Naples, Italy

Correspondence to
Dr Mirko Manetti, Department of Experimental and Clinical Medicine, Section of Anatomy and Histology, University of Florence, Largo Brambilla 3, Florence 50134, Italy; mirkomanetti@yahoo.it, mirko.manetti@unifi.it

Received 18 July 2016
Revised 15 November 2016
Accepted 17 December 2016
Published Online First 6 January 2017

ABSTRACT

Objective Systemic sclerosis (SSc) features multiorgan fibrosis orchestrated predominantly by activated myofibroblasts. Endothelial-to-mesenchymal transition (EndoMT) is a transdifferentiation by which endothelial cells (ECs) lose their specific morphology/markers and acquire myofibroblast-like features. Here, we determined the possible contribution of EndoMT to the pathogenesis of dermal fibrosis in SSc and two mouse models.

Methods Skin sections were immunostained for endothelial CD31 or vascular endothelial (VE)-cadherin in combination with α-smooth muscle actin (α-SMA) myofibroblast marker. Dermal microvascular ECs (dMVECs) were prepared from SSc and healthy skin (SSc-dMVECs and H-dMVECs). H-dMVECs were treated with transforming growth factor-β1 (TGFβ1) or SSc and healthy sera. Endothelial/mesenchymal markers were assessed by real-time PCR, immunoblotting and immunofluorescence. Cell contractile phenotype was assayed by collagen gel contraction.

Results Cells in intermediate stages of EndoMT were identified in dermal vessels of either patients with SSc or bleomycin-induced and urokinase-type plasminogen activator receptor (uPAR)-deficient mouse models. At variance with H-dMVECs, SSc-dMVECs exhibited a spindle-shaped appearance, co-expression of lower levels of CD31 and VE-cadherin with myofibroblast markers (α-SMA+ stress fibres, S100A4 and type I collagen), constitutive nuclear localisation of the EndoMT driver Snail1 and an ability to effectively contract collagen gels. Treatment of H-dMVECs either with SSc sera or TGFβ1 resulted in the acquisition of a myofibroblast-like morphology and contractile phenotype and downregulation of endothelial markers in parallel with the induction of mesenchymal markers. Matrix metalloproteinase-12-dependent uPAR cleavage was implicated in the induction of EndoMT by SSc sera.

Conclusions In SSc, EndoMT may be a crucial event linking endothelial dysfunction and development of dermal fibrosis.

To cite: Manetti M, Romano E, Rosa I, *et al*. *Ann Rheum Dis* 2017;**76**:924–934.

INTRODUCTION

Systemic sclerosis (SSc) is a complex connective tissue disease of unknown aetiology characterised by widespread peripheral microvascular injury evolving into progressive fibrosis of skin and multiple internal organs.[1–3] In SSc, fibrosis results from an unrestrained tissue repair process orchestrated predominantly by activated myofibroblasts that are a population of mesenchymal cells displaying unique biological functions. These include an increased synthesis of fibrillar type I and III collagens, a reduction in the expression of genes encoding extracellular matrix (ECM)-degrading enzymes and α-smooth muscle actin (α-SMA) expression and incorporation into stress fibres, which provides an increased contractile force that is crucial for their tissue remodelling properties.[4–6] Indeed, myofibroblast contraction contributes to a large extent to a progressive increase in connective tissue stiffness, a recently recognised potent profibrotic stimulus.[7–10]

Given the crucial role of myofibroblasts in the pathogenesis of organ fibrosis in a variety of disorders, considerable attention has been paid to the identification of their putative cellular origins. Hence, extensive investigations have revealed that profibrotic myofibroblasts may arise from different sources including expansion and activation of resident tissue fibroblasts and perivascular pericytes, recruitment of bone marrow-derived circulating precursors, transformation of white adipocytes and transdifferentiation of epithelial cells into mesenchymal cells.[4 11–13] More recently, it has been reported with increasing frequency that vascular endothelial cells (ECs) may also exhibit substantial plasticity by undergoing endothelial-to-mesenchymal transition (EndoMT), a transdifferentiation by which ECs disaggregate, lose polarity and acquire ECM-producing myofibroblast features.[14–16] EndoMT is a phenotypical conversion in which ECs downregulate the expression of their specific markers, such as CD31/platelet-EC adhesion molecule-1, von Willebrand factor (vWF) and vascular endothelial (VE)-cadherin, and acquire mesenchymal cell products including α-SMA, S100A4/fibroblast-specific protein-1 (FSP1) and type I collagen, together with stabilisation and nuclear translocation of the transcriptional regulator Snail1, a crucial trigger of mesenchymal transition.[14–16]

To date, EndoMT has emerged as a player in the pathogenesis of tissue fibrosis and fibroproliferative vasculopathy in various diseases, including diabetic nephropathy, cardiac fibrosis, inflammatory bowel disease-related intestinal fibrosis, portal hypertension and primary pulmonary arterial hypertension (PAH).[14 16–21] Of note, extensive

research studies have shown that multiple pathways implicated in SSc pathogenesis, such as transforming growth factor-β (TGFβ), endothelin-1 (ET-1), notch, sonic hedgehog and Wnt pathways, as well as other putative pathways such as oxidative stress and hypoxia, may participate in the molecular mechanisms of the EndoMT process.[16] For instance, EndoMT can be fully induced by TGFβ in cultured ECs from different tissues.[20 22–24]

Although recent studies support the notion that EndoMT may participate in the development of SSc-associated interstitial lung disease (ILD) and PAH,[25 26] the occurrence of such a phenotypical change from ECs to activated myofibroblasts has never been demonstrated in the affected skin of patients with SSc. Therefore, in the present study we combined ex vivo, in vitro and in vivo approaches to investigate the possible contribution of EndoMT to the pathogenesis of dermal fibrosis in SSc and two mouse models of the disease.

MATERIALS AND METHODS

An extended methods section is provided in the online supplementary material.

Cell culture and reagents

Primary cultures of dermal microvascular ECs (dMVECs) were established by explantation from biopsies of the lesional forearm skin from six patients with early diffuse cutaneous SSc (dcSSc; disease duration <2 years from first non-Raynaud symptom)[27] and from six healthy adult subjects under protocols approved by the local institutional review board at the Azienda Ospedaliero-Universitaria Careggi (AOUC), Florence, Italy. Skin biopsies were processed as previously described.[28 29] Patient characteristics are summarised in online supplementary table S1. Adherent cells were detached and subjected to CD31 immunomagnetic isolation by incubation with anti-CD31 conjugated-microbeads.[28 29] Isolated cells were further identified as ECs by labelling with anti-factor VIII-related antigen (vWF) and anti-CD105, followed by reprobing with anti-CD31 antibodies (see online supplementary figure S1). dMVECs from healthy subjects (H-dMVECs) and patients with SSc (SSc-dMVECs) were maintained as detailed in the online supplementary material. In selected experiments, H-dMVECs were treated with recombinant human TGFβ1 (10 ng/mL; PeproTech, Rocky Hill, New Jersey, USA) or 10% serum from patients with early dcSSc (n=6) and healthy subjects (n=6) for 24, 48 and 72 hours. In some experimental points, sera were preincubated with the matrix metalloproteinase-12 (MMP-12) specific inhibitor MMP408 (10 nM; Sigma-Aldrich, St. Louis, Missouri, USA) before cell stimulation.

Fluorescence immunocytochemistry

At the end of the experiments, cells were fixed with 3.7% buffered paraformaldehyde and immunofluorescence with antibodies against CD31, VE-cadherin, α-SMA, S100A4/FSP1, type I collagen and Snail1 (all from Abcam, Cambridge, UK) was performed as detailed in the online supplementary material. In some specimens, Alexa 488-labelled phalloidin (Invitrogen, Carlsbad, California, USA) was applied to the cells to visualise the arrangement of the F-actin cytoskeleton. For primary and secondary antibodies, refer to the online supplementary material.

RNA isolation and quantitative real-time PCR

At the end of the experiments, cultures were harvested and total RNA was isolated using the RNeasy Micro Kit (Qiagen, Milan, Italy). First-strand cDNA synthesis and mRNA quantification by SYBR Green real-time PCR were performed as reported elsewhere.[29] For predesigned oligonucleotide primer pairs, refer to the online supplementary material.

Immunoblotting

Whole-cell protein lysates from dMVECs were subjected to immunoblot analysis as described elsewhere.[29] For details on primary antibodies against CD31, VE-cadherin, α-SMA, S100A4/FSP1, type I collagen, Snail1, Friend leukaemia integration-1 (Fli1), urokinase-type plasminogen activator receptor (uPAR) domain 1 (D1) and domain 2 and α-tubulin, refer to the online supplementary material.

Collagen gel contraction assay

Collagen gel contraction assays were performed as described in the online supplementary material.

ELISA

Levels of MMP-12 in serum samples were measured by quantitative ELISA as described in the online supplementary material.

Fluorescence immunohistochemistry on human and mouse skin

Paraffin-embedded sections of lesional forearm skin biopsies were obtained from 12 patients with SSc (n=4 with limited cutaneous SSc and n=8 with dcSSc) and 10 age-matched and gender-matched healthy donors, as described elsewhere.[28–30] Skin sections from two mouse models of dermal fibrosis were also used. First, 6-week-old male C57BL/6 mice (Charles River Laboratories, Calco, Lecco, Italy) received subcutaneous injections of 100 µL of bleomycin dissolved in 0.9% NaCl (saline solution) at a concentration of 0.5 mg/mL every other day for 4 weeks in well-defined areas of the upper back. Subcutaneous injections of 0.9% NaCl served as controls.[31] The second model consisted of 12-week-old male uPAR-deficient mice and wild-type littermates as described elsewhere.[32 33] All animal protocols were performed in accordance with DL 116/92 and approved by the Institutional Animal Care and Use Committee of the University of Florence. Each experimental group consisted of at least six mice. Double-label immunofluorescence using antibodies against α-SMA and CD31 or VE-cadherin was carried out as detailed in the online supplementary material. The percentage of dermal vessels displaying CD31/α-SMA and VE-cadherin/α-SMA co-localisation was determined in five randomly selected high-power fields of the dermis from each of three sections per sample.

Transmission electron microscopy

Ultrathin skin sections from five patients with dcSSc and five healthy controls were processed and examined according to previously published protocols[34] as detailed in the online supplementary material.

Statistical analysis

Statistical analyses were performed using the Statistical Package for Social Sciences software for Windows, V.20.0 (SPSS, Chicago, Illinois, USA). Data are expressed as means and SEM. The Student's t-test was used for statistical evaluation of the differences between two independent groups. A p value of <0.05 according to a two-tailed distribution was considered statistically significant.

RESULTS

EndoMT in dermal vessels of patients with SSc and experimental models of SSc

In order to determine ex vivo the presence of transitional EndoMT cells, skin sections from patients with SSc and healthy donors were subjected to double immunofluorescence staining for the EC markers CD31 or VE-cadherin and the myofibroblast marker α-SMA. In the healthy dermal microvasculature, α-SMA expression was mostly restricted to pericytes and vascular smooth muscle cells surrounding the endothelial layer (figure 1A). On the contrary, we observed co-localised CD31/α-SMA and VE-cadherin/α-SMA in the endothelium of numerous dermal capillary vessels and arterioles from patients with SSc, suggestive for cells in intermediate stages of EndoMT (figure 1A). Indeed, the percentage of vessels displaying CD31/α-SMA and VE-cadherin/α-SMA co-localisation was significantly increased in skin biopsies from patients with SSc compared with healthy skin ($p<0.001$ for both) (figure 1B). No difference in the frequency of transitional EndoMT cells was observed between SSc cutaneous subsets (data not shown). Furthermore, transmission electron microscopy analysis revealed that the presence of vWF-storing Weibel-Palade bodies was clearly reduced in SSc dermal endothelium (figure 1C).

Next we investigated in vivo the presence of transitional EndoMT cells in the skin of two mouse models of SSc, namely mice with bleomycin-induced dermal fibrosis and uPAR-deficient mice.[31–33] The frequency of transitional EndoMT cells in murine skin was assessed by co-localisation of either CD31 or VE-cadherin and α-SMA. As displayed in figure 2, using both marker combinations we observed transitional EndoMT cells to be present at very low levels in saline-treated control mice, with significantly higher levels in the bleomycin treatment group ($p<0.001$ for both). Similarly, a significantly higher percentage of vessels with CD31/α-SMA and VE-cadherin/α-SMA double-positive cells was detected in the dermis of uPAR-deficient mice compared with wild-type littermates ($p<0.001$ for both) (figure 2A, B).

Cultured SSc-dMVECs co-express endothelial and mesenchymal cell markers and exhibit a myofibroblast-like functional phenotype

The expression of endothelial and mesenchymal cell markers in dMVECs isolated from forearm skin biopsies was investigated by immunofluorescence and immunoblotting. In agreement with previous reports,[28 35] H-dMVECs exhibited a typical endothelial morphology with a polygonal shape, whereas the majority of SSc-dMVECs had an elongated shape often characterised by branches (figure 3A). Both H-dMVECs and SSc-dMVECs were immunopositive for the pan-EC marker CD31 (figure 3A). However, the expression of CD31 and VE-cadherin was markedly decreased in SSc-dMVECs compared with H-dMVECs (figure 3A). SSc-dMVECs also expressed α-SMA, which often was incorporated into stress fibres, as well as S100A4/FSP1 and type I collagen (figure 3A). On the contrary, as expected, in H-dMVECs there was no evidence of α-SMA and type I collagen expression, and S100A4/FSP1 was almost undetectable (figure 3A). Double immunofluorescence staining clearly revealed the unique presence of numerous CD31+ cells displaying α-SMA+ stress fibres in SSc-dMVEC cultures compared with H-dMVECs ($p<0.001$) (figure 3B). Phalloidin staining further revealed that while H-dMVECs showed a weak and disorganised expression of F-actin fibres, SSc-dMVECs exhibited a marked increase in stress fibres mainly organised longitudinally (figure 3A). Furthermore, we investigated the expression of Snail1, a zinc-finger transcription factor that induces numerous transcriptional events leading to the acquisition of a mesenchymal cell-specific phenotype such as stimulation of α-SMA expression.[16 24] As displayed in figure 3A, strong expression and nuclear localisation of Snail1 were constitutively detected in SSc-dMVECs, while Snail1 expression was negligible in H-dMVECs. Immunoblot analyses confirmed either a significantly lower protein expression of CD31 and VE-cadherin or a significantly higher expression of α-SMA, S100A4/FSP1, type I collagen and Snail1 in SSc-dMVECs compared with H-dMVECs ($p<0.001$ for all comparisons) (figure 3C). According to the immunofluorescence data, both α-SMA and type I collagen were undetectable in protein lysates from H-dMVECs (figure 3C). Moreover, SSc-dMVECs exhibited a significant reduction in protein expression of Fli1 ($p<0.001$ vs H-dMVECs) (figure 3C), a transcription factor that plays a pivotal role in the maintenance of EC homeostasis and whose deficiency may be implicated in EndoMT.[36–38] The occurrence of EndoMT was confirmed functionally by the evidence that SSc-dMVECs were able to effectively contract collagen gels (figure 3D).

Treatment with SSc sera induces a myofibroblast-like phenotype in H-dMVECs

Previous studies have demonstrated that treatment with sera from patients with SSc impairs the angiogenic performance of H-dMVECs in vitro.[29 39 40] Nevertheless, whether these anti-angiogenic effects may be in part related to the induction of the EndoMT process has never been investigated. To address this issue, H-dMVECs were challenged with sera from patients with early dcSSc and healthy subjects and subsequently assayed for changes in cell morphology and the expression of endothelial and mesenchymal cell markers. According to the literature,[20 22–24] stimulation with recombinant human TGFβ1 was performed in parallel as a positive control of EndoMT. After 48-hour treatment with SSc sera, H-dMVECs started to disaggregate losing their characteristic polygonal cobblestone-like morphology (figure 4A). These changes progressed rapidly with the appearance of numerous cells exhibiting a spindle-shaped morphology in H-dMVEC cultures treated with SSc sera for 72 hours (figure 4A). As expected, similar findings were observed when H-dMVECs were challenged with TGFβ1, whereas H-dMVEC morphology did not change over time in cultures treated with healthy sera (figure 4A). Indeed, 72-hour treatment either with SSc sera or TGFβ1 induced a significant increase in the percentage of spindle-shaped cells (both $p<0.001$ vs basal H-dMVECs) (figure 4A), which were able to effectively contract collagen gels (figure 4B).

As displayed in figure 5, real-time PCR analysis revealed a significant reduction in mRNA levels of *CD31*, *CDH5* and *FLI1* genes in H-dMVECs treated either with SSc sera or TGFβ1 for 48 hours (all $p<0.001$ vs basal H-dMVECs). This happened in parallel with the induction of *ACTA2*, *S100A4*, *SNAI1*, *COL1A1* and *COL1A2* mRNA expression (all $p<0.001$ vs basal H-dMVECs) (figure 5). On the contrary, 48-hour treatment of H-dMVECs with healthy sera did not affect mRNA expression levels of the aforementioned markers (figure 5). These results were confirmed by immunoblot and immunofluorescence assessment of endothelial and mesenchymal protein expression levels in cells treated for 72 hours (figure 6A–G). In particular, both untreated cells and those treated with healthy sera showed no expression of α-SMA and type I collagen along with very low levels of Snail1, whereas treatment either with SSc sera or TGFβ1 induced the appearance of α-SMA+ stress fibres, de

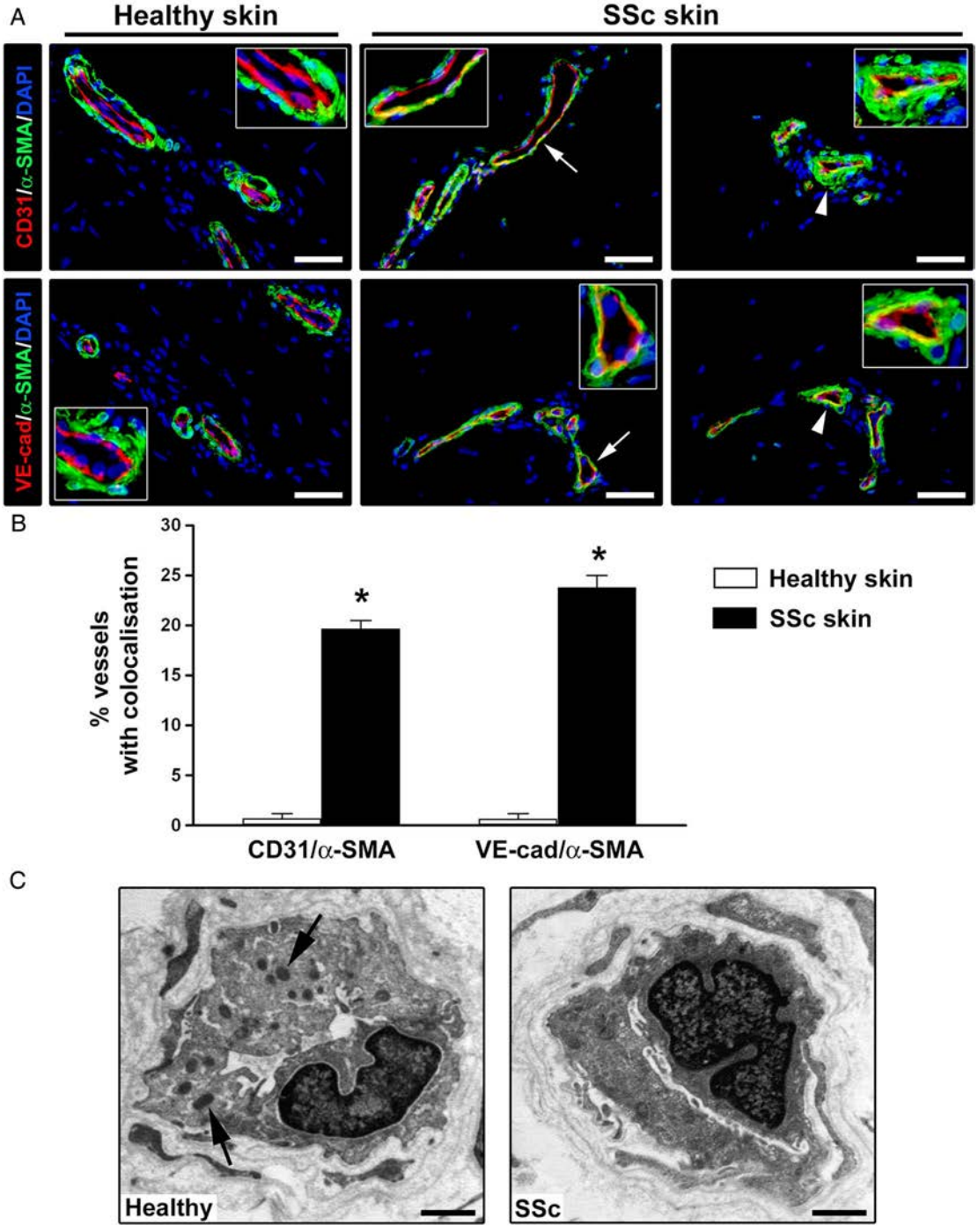


Figure 1 Detection of endothelial-to-mesenchymal transition (EndoMT) in dermal vessels of patients with systemic sclerosis (SSc). (A) Representative fluorescence microphotographs of skin sections from healthy controls and patients with SSc double immunostained for the endothelial cell (EC) markers CD31 or vascular endothelial (VE)-cadherin (red) and the myofibroblast marker α-smooth muscle actin (α-SMA; green) and counterstained with 4′,6-diamidino-2-phenylindole (DAPI; blue) for nuclei. In healthy dermal vessels, α-SMA expression is mostly restricted to pericytes and vascular smooth muscle cells surrounding ECs. In SSc skin, co-localised CD31/α-SMA and VE-cadherin/α-SMA give rise to yellow staining, which is evident in transitional EndoMT cells of numerous capillary vessels (arrows) and arterioles (arrowheads). In each panel, insets show higher magnification views of dermal microvessels. Scale bar=50 μm. (B) The percentage of dermal vessels displaying CD31/α-SMA and VE-cadherin/α-SMA co-localisation is significantly increased in skin biopsies from patients with SSc (n=12) compared with healthy skin (n=10). Data are mean ±SEM. *p<0.001 versus healthy skin. (C) Representative transmission electron microscopy microphotographs of dermal capillary vessels from healthy controls (n=5) and patients with SSc (n=5). At least eight capillary vessels from each of three ultrathin sections per sample were analysed. Numerous Weibel-Palade bodies (arrows) are present in healthy dermal ECs, while they are reduced or even absent in SSc dermal ECs. Scale bar=2 μm.

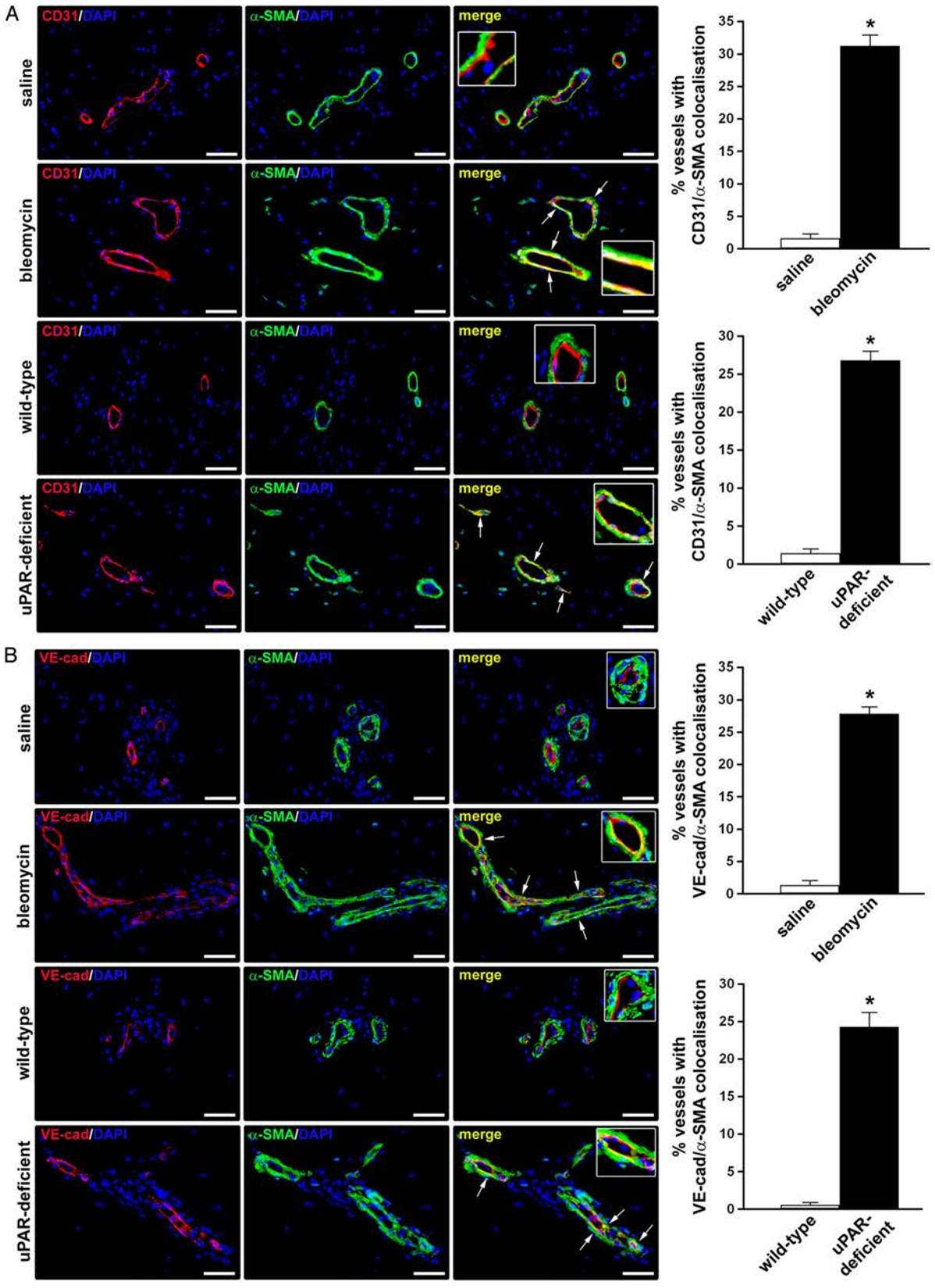


Figure 2 Detection of endothelial-to-mesenchymal transition (EndoMT) in dermal vessels of murine models of systemic sclerosis (SSc). (A, B) Representative fluorescence microphotographs of mouse skin sections double immunostained for either CD31 (red) (A) or vascular endothelial (VE)-cadherin (red) (B) endothelial cell markers and the myofibroblast marker α-smooth muscle actin (α-SMA; green) with 4′,6-diamidino-2-phenylindole (DAPI; blue) counterstain for nuclei are shown. In the dermis of bleomycin-treated mice and urokinase-type plasminogen activator receptor (uPAR)-deficient mice, co-localisation of either CD31 or VE-cadherin and α-SMA gives rise to yellow staining, which is evident in transitional EndoMT cells of numerous microvessels (arrows). Insets show higher magnification views of dermal microvessels from the corresponding panels. Scale bar=50 μm. The percentage of dermal vessels displaying CD31/α-SMA or VE-cadherin/α-SMA co-localisation is reported in the histograms. Data are mean±SEM (six mice in each experimental group). *p<0.001 versus saline-treated mice (A, B, top), *p<0.001 versus wild-type littermates (A, B, bottom).

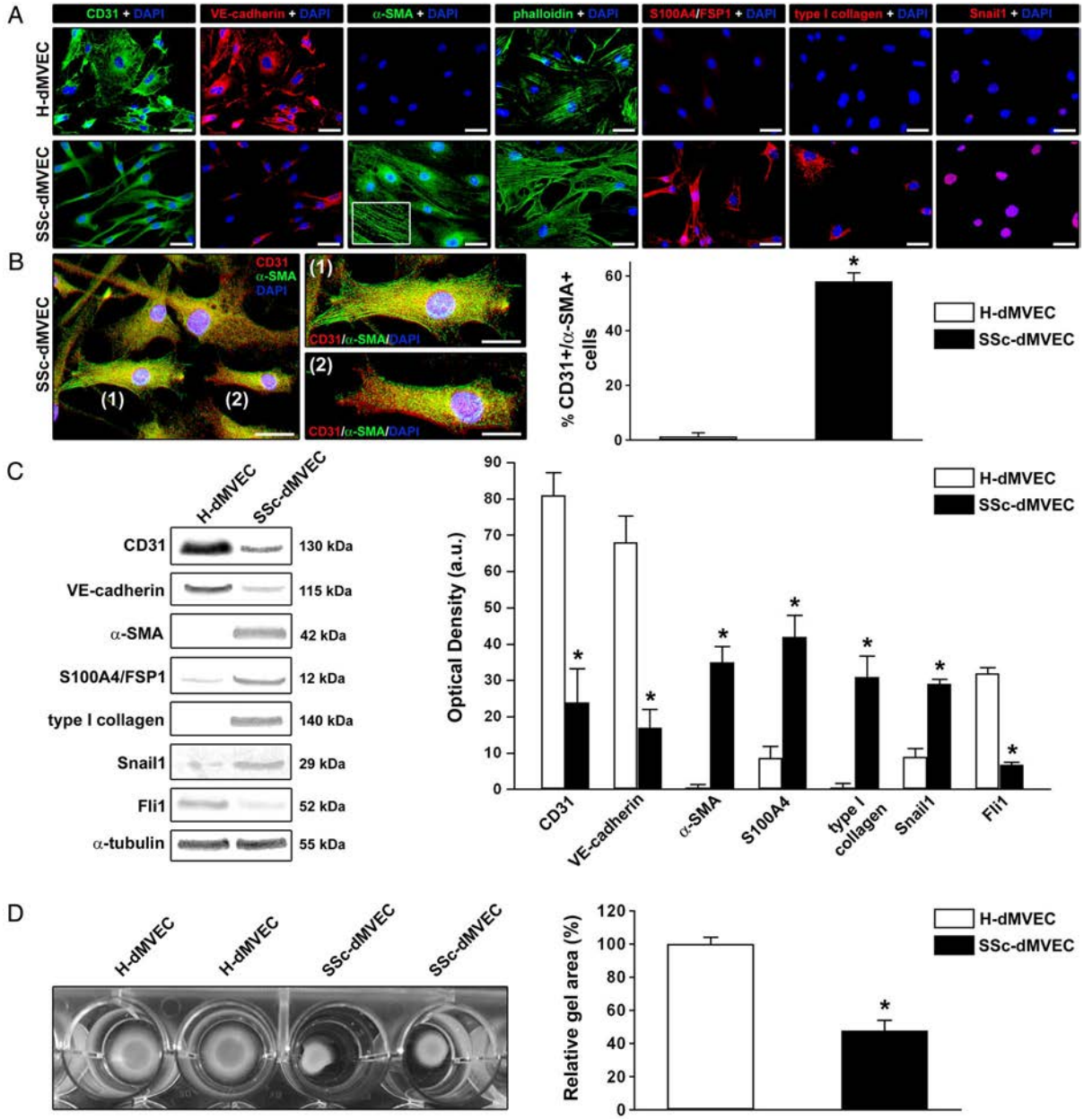


Figure 3 Dermal microvascular endothelial cells (dMVECs) isolated from systemic sclerosis (SSc) skin co-express endothelial and mesenchymal cell markers and exhibit a myofibroblast-like functional phenotype. (A) Representative fluorescence microphotographs of healthy dMVECs (H-dMVECs) and SSc-dMVECs (n=6 each) immunostained for CD31, vascular endothelial (VE)-cadherin, α-smooth muscle actin (α-SMA), F-actin (phalloidin), S100A4/fibroblast-specific protein-1 (FSP1), type I collagen and Snail1 transcription factor. Nuclei are counterstained with 4′,6-diamidino-2-phenylindole (DAPI). Both H-dMVECs and SSc-dMVECs are immunopositive for the pan-endothelial cell marker CD31. The expression of CD31 and VE-cadherin is markedly lower in SSc-dMVECs compared with H-dMVECs. SSc-dMVECs exhibit α-SMA+ stress fibres (shown at higher magnification in the inset), a marked increase in phalloidin+ stress fibres mainly organised longitudinally, and expression of S100A4/FSP1, type I collagen and nuclear Snail1. In H-dMVECs, α-SMA and type I collagen are undetectable, while expression of S100A4/FSP1 and Snail1 is negligible. Scale bar=50 μm. (B) Representative fluorescence microphotographs of SSc-dMVECs double immunostained for CD31 (red) and α-SMA (green) with DAPI (blue) counterstain for nuclei. Note the presence of CD31+ cells displaying α-SMA+ stress fibres. Cells labelled as (1) and (2) in the left panel are shown at higher magnification in the right panels. The degree of α-SMA arrangement into stress fibres varies among cells. Scale bar=50 μm (left panel), 20 μm (right panels). The percentage of CD31/α-SMA double-positive cells is reported in the histograms. Data are mean±SEM. *p<0.001 versus H-dMVECs. (C) Protein lysates from H-dMVECs and SSc-dMVECs were assayed for the expression of CD31, VE-cadherin, α-SMA, S100A4/FSP1, type I collagen, Snail1 and Friend leukaemia integration-1 (Fli1). Representative immunoblots are shown. Molecular weight values (kDa) are indicated. The densitometric analysis of the bands normalised to α-tubulin is reported in the histograms. Data are mean±SEM of optical density in arbitrary units (a.u.). *p<0.001 versus H-dMVECs. Results are representative of three independent experiments performed with each of the six H-dMVEC and SSc-dMVEC lines. (D) Collagen gel contraction assay with H-dMVECs and SSc-dMVECs (n=6 each). Gel size in the presence of SSc-dMVECs is expressed as percentage of that observed in the presence of H-dMVECs. Data are mean±SEM. *p<0.001 versus H-dMVECs.

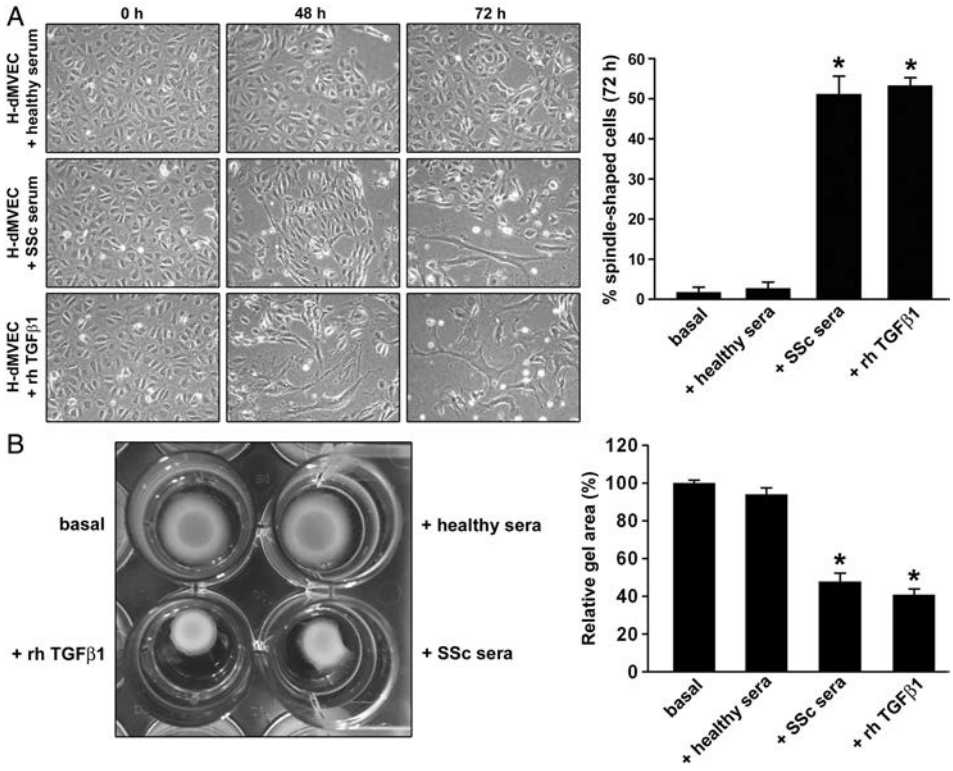


Figure 4 Treatment with sera from patients with systemic sclerosis (SSc) induces a myofibroblast-like morphology and functional phenotype in healthy dermal microvascular endothelial cells (H-dMVECs). (A) Representative phase-contrast microphotographs of H-dMVECs (n=3) at baseline and after treatment for 48 and 72 hours with sera from healthy subjects (n=6), sera from patients with SSc (n=6) or recombinant human transforming growth factor-β1 (rh TGFβ1; 10 ng/mL) are shown (×10 original magnification). The morphology of H-dMVECs does not change over time in cultures treated with healthy sera. After 48-hour treatment either with SSc sera or rh TGFβ1, H-dMVECs start to disaggregate and lose their characteristic polygonal cobblestone-like morphology. Cells exhibiting a spindle-shaped morphology are clearly visible in H-dMVEC cultures treated either with SSc sera or rh TGFβ1 for 72 hours. The percentage of spindle-shaped cells is reported in the histograms. Data are mean±SEM. *p<0.001 versus basal H-dMVECs. (B) Collagen gel contraction assay with H-dMVECs at baseline and after treatment for 72 hours with healthy sera (n=6), SSc sera (n=6) or rh TGFβ1. Gel size in the different experimental conditions is expressed as percentage of baseline. Data are mean±SEM. *p<0.001 versus basal H-dMVECs.

novo synthesis of type I collagen and strong expression and nuclear localisation of Snail1 (figure 6B–G).

MMP-12-dependent cleavage of uPAR is implicated in the induction of EndoMT by SSc sera

We previously demonstrated that in SSc-dMVECs uPAR undergoes a MMP-12-dependent cleavage of domain D1 resulting in impaired angiogenesis.[35 41] Interestingly, the cleavage of uPAR-D1 was shown to be a crucial step in fibroblast-to-myofibroblast transition.[42] Therefore, we herein investigated whether MMP-12-dependent uPAR-D1 cleavage could be implicated in the induction of EndoMT by SSc sera. Consistent with previous reports,[40 43] MMP-12 levels were raised in SSc sera (see online supplementary figure S2A). Treatment of H-dMVECs with SSc sera resulted in uPAR-D1 cleavage already after 24 hours (see online supplementary figure S2B). Such a cleavage was instead prevented when SSc sera were preincubated with the MMP-12-specific inhibitor MMP408 (see online supplementary figure S2B). As shown in online supplementary figure S3, preincubation with MMP408 significantly blunted the effects of 48-hour treatment with SSc sera on gene expression of endothelial and mesenchymal cell markers.

DISCUSSION

Our data provide the first direct evidence that EndoMT may take place in the skin of patients with SSc and may have therefore a role in the pathogenesis of dermal fibrosis. The ex vivo immunohistological data clearly demonstrate the presence of transitional EndoMT cells simultaneously expressing EC and myofibroblast markers in SSc dermal microvasculature. In contrast, EndoMT was only observed at negligible levels in control skin. These results are substantially in agreement with similar findings recently described in the pulmonary vessels of patients with SSc-associated ILD and PAH.[25 26] We have further characterised in vitro the phenotype of dMVECs isolated from SSc skin and found that these cells are in an intermediate state between an EC and a myofibroblast-like contractile phenotype, combining markers of both cell types. The results also show that H-dMVECs can undergo EndoMT in response to treatment with SSc sera, thus supporting the hypothesis that such cellular transdifferentiation may be operative in SSc. In fact, after a prolonged challenge with SSc sera, H-dMVECs lost their typical endothelial cobblestone appearance and acquired myofibroblast-like structural and functional features. Consistent with these morphofunctional changes, SSc serum-treated H-dMVECs exhibited a reduction in the expression of EC markers CD31 and VE-cadherin and an upregulation of mesenchymal markers, including α-SMA+ stress fibres, S100A4/FSP1, type I collagen and nuclear Snail1. Furthermore, the presence of transitional EndoMT cells in dermal vessels of two murine models of SSc is a matter of interest. Indeed, previous studies have demonstrated the

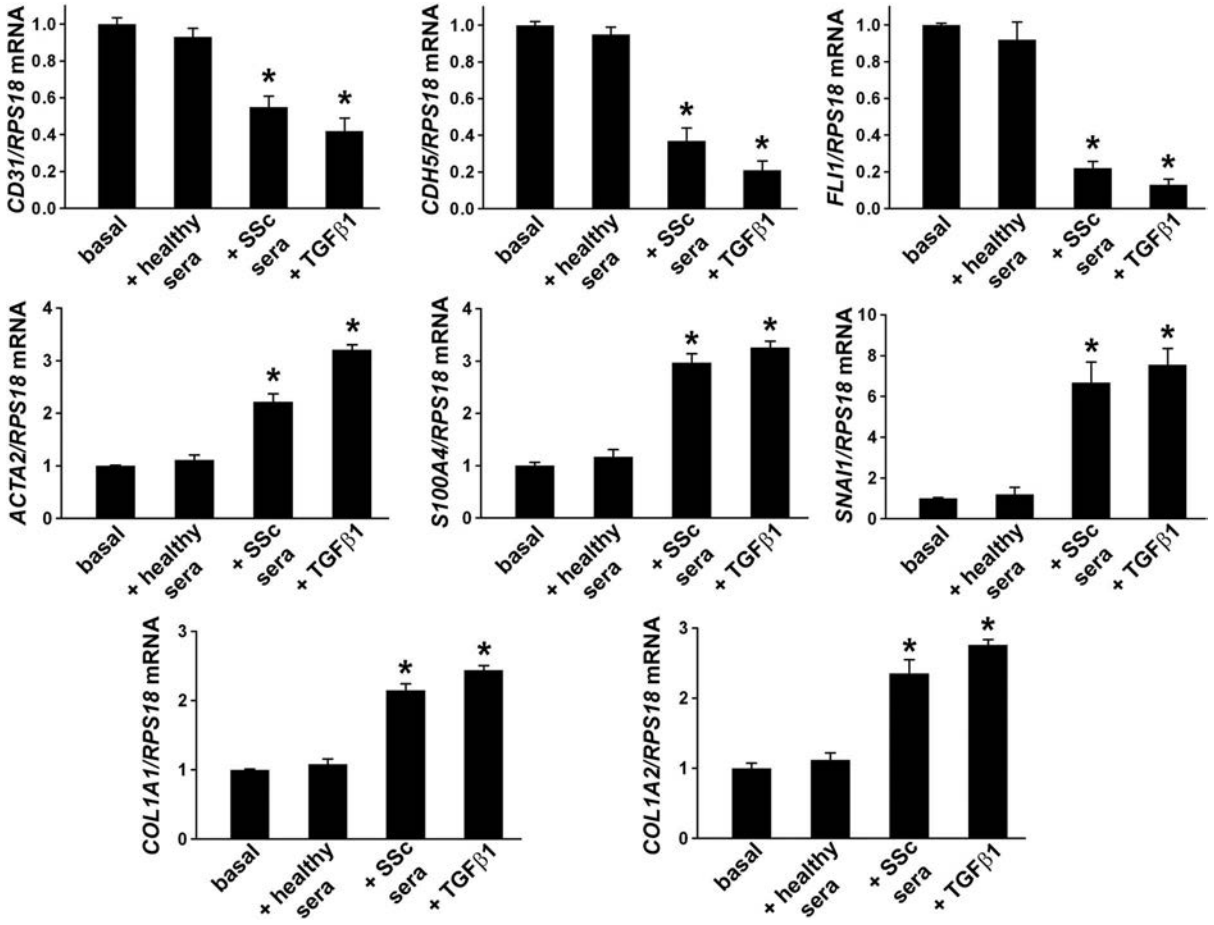


Figure 5 Treatment with sera from patients with systemic sclerosis (SSc) induces changes in mRNA expression levels of endothelial and mesenchymal cell markers in healthy dermal microvascular endothelial cells (H-dMVECs). H-dMVECs were treated for 48 hours with sera from healthy subjects (n=6), sera from patients with SSc (n=6) or recombinant human transforming growth factor-β1 (TGFβ1; 10 ng/mL) and subsequently assayed for mRNA expression levels of *CD31*, *CDH5* (vascular endothelial (VE)-cadherin), *FLI1*, *ACTA2* (α-smooth muscle actin (α-SMA)), *S100A4*, *SNAI1* (Snail1), *COL1A1* and *COL1A2* genes by quantitative real-time PCR. Ribosomal protein S18 (*RPS18*) mRNA was measured as an endogenous control for normalisation. The relative values compared with basal H-dMVECs are expressed as mean±SEM of three independent experiments performed with three H-dMVEC lines. *p<0.001 versus basal H-dMVECs.

occurrence of EndoMT in animal models of cardiac, pulmonary and renal fibrosis, as well as in models of PAH.[16 23 26 44 45] Although our experimental data support the notion that EndoMT may contribute to the accumulation of myofibroblasts and the development of dermal fibrosis in vivo, this needs to be further confirmed by using lineage tracing in different preclinical models of SSc.

Besides the increase in the number of profibrotic myofibroblasts, EndoMT may favour microvascular derangement and loss of ECs contributing to capillary rarefaction, impaired angiogenesis and chronic tissue ischaemia in SSc skin. Indeed, endothelial dysfunction is considered a pivotal factor contributing to peripheral vessel remodelling in SSc.[3 15 41] A defective response to proangiogenic stimuli and several functional defects, such as an impaired ability to organise into capillary-like tubes in vitro, have been extensively reported in SSc-dMVECs.[28 29 35 41 46] Moreover, transcriptome profiling studies have revealed profound differences in the expression of genes encoding a variety of angiogenic regulators between SSc-dMVECs and H-dMVECs.[41 47] In this context, our present findings shed light on EndoMT as a pathogenic mechanism that in SSc may directly link EC dysfunction to the development of dermal fibrosis. The intrinsic propensity of SSc-dMVECs to transition towards a profibrotic myofibroblast-like phenotype might in effect largely explain their well-known defective angiogenic behaviour. In addition, here we clearly demonstrate that a prolonged treatment with sera from patients with SSc is capable of sustaining the EndoMT process in H-dMVECs. Of note, shorter time treatments with SSc sera have previously been shown to impair angiogenesis and survival of H-dMVECs.[29 39 40] Mechanistically, our present findings show that MMP-12-dependent cleavage of uPAR, a process that has been deeply implicated either in the impaired angiogenic performance of SSc-dMVECs or in fibroblast-to-myofibroblast differentiation,[35 41 42] takes part in the pro-EndoMT effect exerted by SSc sera. Besides MMP-12, additional as yet unidentified circulating factors might trigger EndoMT and the loss of microvascular integrity in SSc dermis. Though further in-depth studies will be required, potential candidates include a large array of mediators that are elevated in SSc and have been demonstrated to induce EndoMT in vitro, such as TGFβ1, ET-1, tumour necrosis factor-α, asymmetric dimethylarginine and endostatin.[16 19 26 48 49] Consistent with our in vitro observations, a recent study reported that sera from patients with chronic kidney disease induced EndoMT, decreased proliferation and increased apoptosis of human coronary artery ECs.[49]

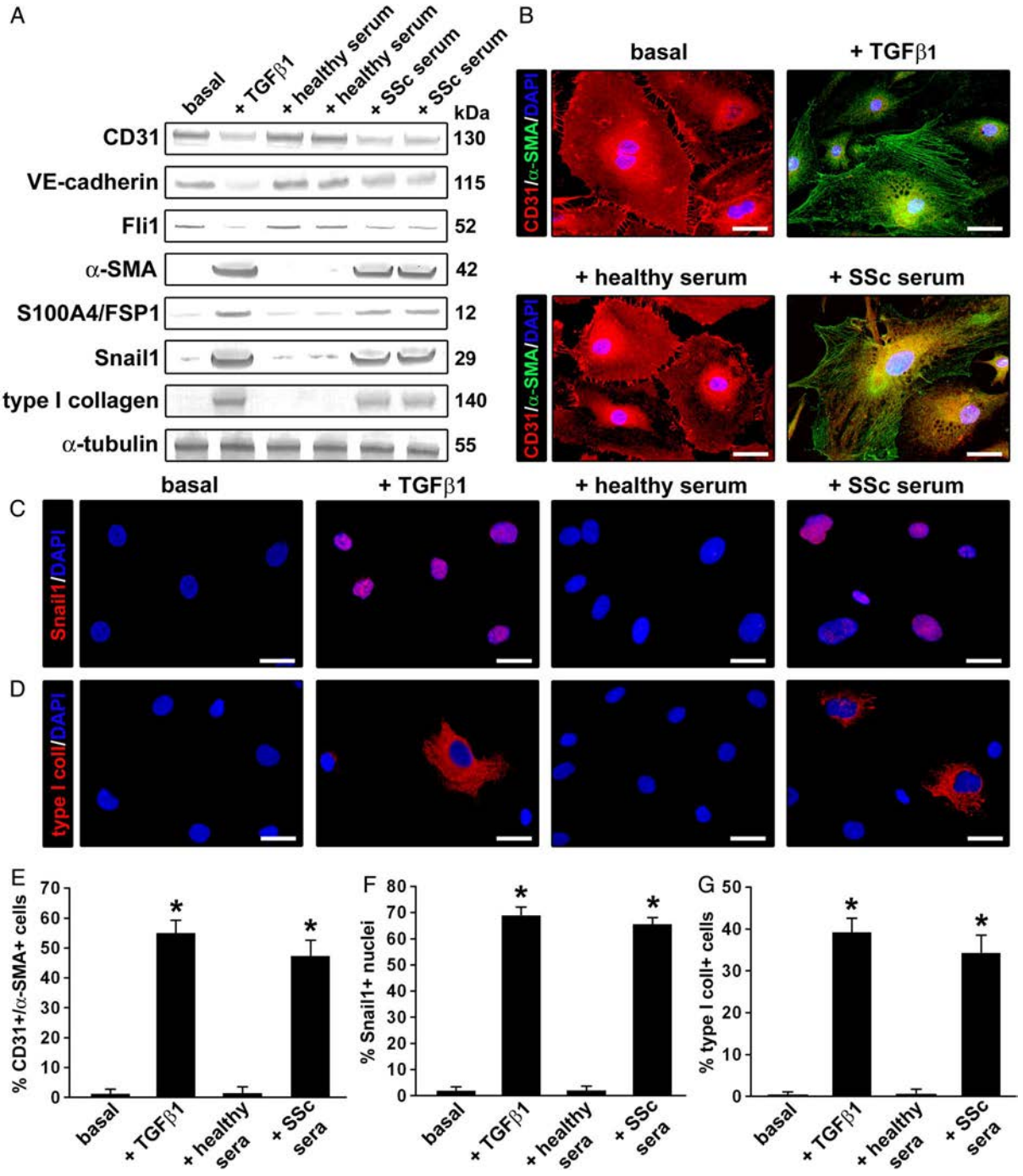


Figure 6 Treatment with sera from patients with systemic sclerosis (SSc) induces changes in protein expression levels of endothelial and mesenchymal cell markers in healthy dermal microvascular endothelial cells (H-dMVECs). H-dMVECs were treated for 72 hours with sera from healthy subjects (n=6), sera from patients with SSc (n=6) or recombinant human transforming growth factor-β1 (TGFβ1; 10 ng/mL) and subsequently assayed for protein expression levels of CD31, vascular endothelial (VE)-cadherin, Friend leukaemia integration-1 (Fli1), α-smooth muscle actin (α-SMA), S100A4/fibroblast-specific protein-1 (FSP1), Snail1 and type I collagen. (A) Representative immunoblots are shown. Molecular weight values (kDa) are indicated. Protein expression of α-tubulin was measured as a loading control. Results are representative of three independent experiments performed with three H-dMVEC lines. (B–D) Representative fluorescence microphotographs show H-dMVECs double immunostained for the endothelial cell marker CD31 (red) and the myofibroblast marker α-SMA (green), or immunostained for Snail1 (red) and type I collagen (red). Nuclei are counterstained with 4′,6-diamidino-2-phenylindole (DAPI; blue). Treatment of H-dMVECs either with SSc sera or TGFβ1, but not with healthy sera, induces downregulation of CD31 in parallel with the appearance of α-SMA+ stress fibres, strong expression and nuclear localisation of Snail1 and de novo synthesis of type I collagen. Scale bar=20 μm. (E–G) The percentage of CD31/α-SMA double-positive cells (E), Snail1+ nuclei (F) and type I collagen+ cells (G) is reported in the histograms. Data are mean±SEM. *p<0.001 versus basal H-dMVECs.

These effects were mainly attributable to increased concentrations of circulating angiogenesis and nitric oxide inhibitors.[49] Finally, when considering the autoimmune background of SSc, we cannot exclude the possible implication of functional (agonistic) autoantibodies against cell surface receptors in the EndoMT process. Indeed, a high proportion of patients with SSc display agonistic autoantibodies against the angiotensin II type 1 receptor and the ET-1 type A receptor, which can

induce a variety of cellular responses such as production of TGFβ by dMVECs and synthesis of type I collagen by skin fibroblasts.[50] Further mechanistic studies aimed at identifying key initiators of EndoMT in SSc are warranted.

In summary, our data collectively support the notion that EndoMT is a process occurring in the dermal endothelium of patients with SSc, where it may represent a crucial link between EC dysfunction and development of fibrosis. Hence, preventing or blocking EndoMT might be a novel and useful approach to treat peripheral microvasculopathy and prevent, at least in part, skin fibrosis in patients with SSc.

Contributors Study conception and design: MM, ER, LI-M and MM-C. Acquisition of data: MM, ER, IR, SG, SB-R, ADP, LI-M and MM-C. Interpretation of data: MM, ER, IR, LI-M and MM-C. Manuscript preparation: MM and MM-C.

Funding Supported by grants from the University of Florence (Progetti di Ricerca di Ateneo to LI-M and MM-C).

Competing interests None declared.

Patient consent Obtained.

Ethics approval The study was approved by the local institutional review board at the Azienda Ospedaliero-Universitaria Careggi (AOUC), Florence, Italy.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Allanore Y, Simms R, Distler O, *et al*. Systemic sclerosis. *Nat Rev Dis Primers* 2015;1:15002.
2 Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. *N Engl J Med* 2009;360:1989–2003.
3 Matucci-Cerinic M, Kahaleh B, Wigley FM. Evidence that systemic sclerosis is a vascular disease. *Arthritis Rheum* 2013;65:1953–62.
4 Hinz B, Phan SH, Thannickal VJ, *et al*. Recent developments in myofibroblast biology: paradigms for connective tissue remodeling. *Am J Pathol* 2012;180:1340–55.
5 Hu B, Phan SH. Myofibroblasts. *Curr Opin Rheumatol* 2013;25:71–7.
6 Kis K, Liu X, Hagood JS. Myofibroblast differentiation and survival in fibrotic disease. *Expert Rev Mol Med* 2011;13:e27.
7 Hinz B. The extracellular matrix and transforming growth factor-β1: tale of a strained relationship. *Matrix Biol* 2015;47:54–65.
8 Hinz B. The myofibroblast: paradigm for a mechanically active cell. *J Biomech* 2010;43:146–55.
9 Huang X, Yang N, Fiore VF, *et al*. Matrix stiffness-induced myofibroblast differentiation is mediated by intrinsic mechanotransduction. *Am J Respir Cell Mol Biol* 2012;47:340–8.
10 Gerber EE, Gallo EM, Fontana SC, *et al*. Integrin-modulating therapy prevents fibrosis and autoimmunity in mouse models of scleroderma. *Nature* 2013;503:126–30.
11 Hinz B, Phan SH, Thannickal VJ, *et al*. The myofibroblast: one function, multiple origins. *Am J Pathol* 2007;170:1807–16.
12 Ebmeier S, Horsley V. Origin of fibrosing cells in systemic sclerosis. *Curr Opin Rheumatol* 2015;27:555–62.
13 Marangoni RG, Korman BD, Wei J, *et al*. Myofibroblasts in murine cutaneous fibrosis originate from adiponectin-positive intradermal progenitors. *Arthritis Rheumatol* 2015;67:1062–73.
14 Piera-Velazquez S, Li Z, Jimenez SA. Role of endothelial-mesenchymal transition (EndoMT) in the pathogenesis of fibrotic disorders. *Am J Pathol* 2011;179:1074–80.
15 Manetti M, Guiducci S, Matucci-Cerinic M. The origin of the myofibroblast in fibroproliferative vasculopathy: does the endothelial cell steer the pathophysiology of systemic sclerosis? *Arthritis Rheum* 2011;63:2164–7.
16 Jimenez SA, Piera-Velazquez S. Endothelial to mesenchymal transition (EndoMT) in the pathogenesis of Systemic Sclerosis-associated pulmonary fibrosis and pulmonary arterial hypertension. Myth or reality? *Matrix Biol* 2016;51:26–36.
17 He J, Xu Y, Koya D, *et al*. Role of the endothelial-to-mesenchymal transition in renal fibrosis of chronic kidney disease. *Clin Exp Nephrol* 2013;17:488–97.
18 Li J, Bertram JF. Review: Endothelial-myofibroblast transition, a new player in diabetic renal fibrosis. *Nephrology (Carlton)* 2010;15:507–12.
19 Rieder F, Kessler SP, West GA, *et al*. Inflammation-induced endothelial-to-mesenchymal transition: a novel mechanism of intestinal fibrosis. *Am J Pathol* 2011;179:2660–73.
20 Kitao A, Sato Y, Sawada-Kitamura S, *et al*. Endothelial to mesenchymal transition via transforming growth factor-beta1/Smad activation is associated with portal venous stenosis in idiopathic portal hypertension. *Am J Pathol* 2009;175:616–26.
21 Ranchoux B, Antigny F, Rucker-Martin C, *et al*. Endothelial-to-mesenchymal transition in pulmonary hypertension. *Circulation* 2015;131:1006–18.
22 Arciniegas E, Sutton AB, Allen TD, *et al*. Transforming growth factor beta 1 promotes the differentiation of endothelial cells into smooth muscle-like cells in vitro. *J Cell Sci* 1992;103(Pt 2):521–9.
23 Zeisberg EM, Potenta SE, Sugimoto H, *et al*. Fibroblasts in kidney fibrosis emerge via endothelial-to-mesenchymal transition. *J Am Soc Nephrol* 2008;19:2282–7.
24 Li Z, Jimenez SA. Protein kinase Cδ and c-Abl kinase are required for transforming growth factor β induction of endothelial-mesenchymal transition in vitro. *Arthritis Rheum* 2011;63:2473–83.
25 Mendoza FA, Piera-Velazquez S, Farber JL, *et al*. Endothelial cells expressing endothelial and mesenchymal cell gene products in lung tissue from patients with systemic sclerosis-associated interstitial lung disease. *Arthritis Rheumatol* 2016;68:210–17.
26 Good RB, Gilbane AJ, Trinder SL, *et al*. Endothelial to mesenchymal transition contributes to endothelial dysfunction in pulmonary arterial hypertension. *Am J Pathol* 2015;185:1850–8.
27 van den Hoogen F, Khanna D, Fransen J, *et al*. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis* 2013;72:1747–55.
28 Manetti M, Guiducci S, Romano E, *et al*. Overexpression of VEGF165b, an inhibitory splice variant of vascular endothelial growth factor, leads to insufficient angiogenesis in patients with systemic sclerosis. *Circ Res* 2011;109:e14–26.
29 Romano E, Chora I, Manetti M, *et al*. Decreased expression of neuropilin-1 as a novel key factor contributing to peripheral microvasculopathy and defective angiogenesis in systemic sclerosis. *Ann Rheum Dis* 2016;75:1541–9.
30 Manetti M, Guiducci S, Romano E, *et al*. Decreased expression of the endothelial cell-derived factor EGFL7 in systemic sclerosis: potential contribution to impaired angiogenesis and vasculogenesis. *Arthritis Res Ther* 2013;15:R165.
31 Distler JH, Jüngel A, Huber LC, *et al*. Imatinib mesylate reduces production of extracellular matrix and prevents development of experimental dermal fibrosis. *Arthritis Rheum* 2007;56:311–22.
32 Manetti M, Rosa I, Milia AF, *et al*. Inactivation of urokinase-type plasminogen activator receptor (uPAR) gene induces dermal and pulmonary fibrosis and peripheral microvasculopathy in mice: a new model of experimental scleroderma? *Ann Rheum Dis* 2014;73:1700–9.
33 Manetti M, Rosa I, Fazi M, *et al*. Systemic sclerosis-like histopathological features in the myocardium of uPAR-deficient mice. *Ann Rheum Dis* 2016;75:474–8.
34 Manetti M, Guiducci S, Ruffo M, *et al*. Evidence for progressive reduction and loss of telocytes in the dermal cellular network of systemic sclerosis. *J Cell Mol Med* 2013;17:482–96.
35 Margheri F, Manetti M, Serratì S, *et al*. Domain 1 of the urokinase-type plasminogen activator receptor is required for its morphologic and functional, beta2 integrin-mediated connection with actin cytoskeleton in human microvascular endothelial cells: failure of association in systemic sclerosis endothelial cells. *Arthritis Rheum* 2006;54:3926–38.
36 Asano Y, Stawski L, Hant F, *et al*. Endothelial Fli1 deficiency impairs vascular homeostasis: a role in scleroderma vasculopathy. *Am J Pathol* 2010;176:1983–98.
37 Taniguchi T, Asano Y, Akamata K, *et al*. Fibrosis, vascular activation, and immune abnormalities resembling systemic sclerosis in bleomycin-treated Fli-1-haploinsufficient mice. *Arthritis Rheumatol* 2015;67:517–26.
38 Manetti M. Fli1 deficiency and beyond: a unique pathway linking peripheral vasculopathy and dermal fibrosis in systemic sclerosis. *Exp Dermatol* 2015;24:256–7.
39 Romano E, Bellando-Randone S, Manetti M, *et al*. Bosentan blocks the antiangiogenic effects of sera from systemic sclerosis patients: an in vitro study. *Clin Exp Rheumatol* 2015;33(Suppl 91):S148–52.
40 Borghini A, Manetti M, Nacci F, *et al*. Systemic sclerosis sera impair angiogenic performance of dermal microvascular endothelial cells: therapeutic implications of cyclophosphamide. *PLoS ONE* 2015;10:e0130166.
41 Manetti M, Guiducci S, Ibba-Manneschi L, *et al*. Mechanisms in the loss of capillaries in systemic sclerosis: angiogenesis versus vasculogenesis. *J Cell Mol Med* 2010;14:1241–54.
42 Bernstein AM, Twining SS, Warejcka DJ, *et al*. Urokinase receptor cleavage: a crucial step in fibroblast-to-myofibroblast differentiation. *Mol Biol Cell* 2007;18:2716–27.
43 Manetti M, Guiducci S, Romano E, *et al*. Increased serum levels and tissue expression of matrix metalloproteinase-12 in patients with systemic sclerosis: correlation with severity of skin and pulmonary fibrosis and vascular damage. *Ann Rheum Dis* 2012;71:1064–72.
44 Hashimoto N, Phan SH, Imaizumi K, *et al*. Endothelial-mesenchymal transition in bleomycin-induced pulmonary fibrosis. *Am J Respir Cell Mol Biol* 2010;43:161–72.
45 Zeisberg EM, Tarnavski O, Zeisberg M, *et al*. Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. *Nat Med* 2007;13:952–61.
46 Tsou PS, Rabquer BJ, Ohara RA, *et al*. Scleroderma dermal microvascular endothelial cells exhibit defective response to pro-angiogenic chemokines. *Rheumatology (Oxford)* 2016;55:745–54.

47 Giusti B, Fibbi G, Margheri F, *et al*. A model of anti-angiogenesis: differential transcriptosome profiling of microvascular endothelial cells from diffuse systemic sclerosis patients. *Arthritis Res Ther* 2006;8:R115.
48 Wermuth PJ, Li Z, Mendoza FA, *et al*. Stimulation of transforming growth factor-β1-induced endothelial-to-mesenchymal transition and tissue fibrosis by endothelin-1 (ET-1): a novel profibrotic effect of ET-1. *PLoS ONE* 2016;11:e0161988.
49 Charytan DM, Padera R, Helfand AM, *et al*. Increased concentration of circulating angiogenesis and nitric oxide inhibitors induces endothelial to mesenchymal transition and myocardial fibrosis in patients with chronic kidney disease. *Int J Cardiol* 2014;176:99–109.
50 Cabral-Marques O, Riemekasten G. Vascular hypothesis revisited: Role of stimulating antibodies against angiotensin and endothelin receptors in the pathogenesis of systemic sclerosis. *Autoimmun Rev* 2016;15:690–4.

Criteria for CAPS, is it all in the name?

This paper[1] describes an original work conducted by an International team of 16 recognised clinical experts in the field of autoinflammatory diseases. The aim of this consortium was to develop diagnostic criteria for cryopyrin-associated periodic syndrome (CAPS). They resulted in a model that indisputably is relevant to describe these rare and heterogeneous diseases among other autoinflammatory diseases. Their proposed CAPS diagnosis criteria are primarily clinical.

We would like to comment on the pathophysiological mechanism underlying 'CAPS'. The *NLRP3* gene encodes cryopyrin, the historical name of the NLRP3 protein, a key component of the NLRP3 inflammasome.[2] As clearly stated by the authors, 'cryopyrin-associated periodic syndrome' includes, by definition, a group of diseases associated with *NLRP3* mutations. Deriving from this concept, a 'non-CAPS autoinflammatory disease' would correspond to a disease caused by mutations/polymorphisms in one or several other gene(s). The consortium nevertheless concluded that 'most importantly', the model 'did not mandate evidence of a disease-causing *NLRP3* mutation'. Indeed, somatic mosaicism can be missed by conventional sequencing approaches.[3] However, the proposed criteria cannot stricto sensu be restricted to CAPS as it is likely to encompass syndromes caused by mutations in autoinflammatory genes other than *NLRP3*. For example, familial cold autoinflammatory syndromes (FCAS) can result from a number of recently discovered genes. FCAS1 (online inheritance in men (OMIM) no. 120100) is the classical FCAS caused by mutations in *NLRP3*; FCAS2 (OMIM no. 611762, also known as NLRP12-associated periodic syndrome (NAPS12), NLRP12-associated periodic syndrome), FCAS3 (OMIM no. 614468, also known as PLCG2-associated antibody deficiency and immune dysregulation (PLAID), Phospholipase C Gamma 2 (PLCG2)-associated antibody deficiency and immune dysregulation) and FCAS4 (OMIM no. 616115, also known as AIFEC, autoinflammation with infantile enterocolitis or recurrent macrophage activation syndrome) are due to the mutations in *NLRP12*, *PLCG2* and *NLRC4*, respectively. The proposed criteria (raised inflammatory markers, urticaria-like rash and arthralgia) would perform well with these four FCAS, but only FCAS1 is a CAPS. The authors fairly admitted that 'the number of CAPS cases and controls was limited and not all possible differential diagnoses of CAPS may have been included, potentially leading to an overestimation of the specificity of the proposed model'.

The issue of the disease name, reflecting either the main symptoms or the molecular mechanisms of the condition, has been raised many times in the autoinflammatory diseases community, and propositions for refined taxonomy will shortly emerge. The debate is not just semantic as differential therapeutic approaches can be taken according to the molecular defect in cause in the condition presented by the patient.

Isabelle Touitou, Guillaume Sarrabay

Reference Centre for Autoinflammatory Diseases, Hôpital A de Villeneuve, CHRU Montpellier, Montpellier, France

Correspondence to Professor Isabelle Touitou, Reference Centre for Autoinflammatory Diseases, Hôpital A de Villeneuve, CHRU Montpellier, 34295 Montpellier Cedex 5, France; Isabelle.touitou@inserm.fr

Competing interests None.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Touitou I, Sarrabay G. *Ann Rheum Dis* 2017;**76**:e9.

Received 14 October 2016
Accepted 18 October 2016
Published Online First 16 November 2016

Ann Rheum Dis 2017;**76**:e9. doi:10.1136/annrheumdis-2016-210681

REFERENCES

1 Kuemmerle-Deschner JB, Ozen S, Tyrrell PN, *et al*. Diagnostic criteria for cryopyrin-associated periodic syndrome (CAPS). *Ann Rheum Dis* 2016; Published Online First 4 October 2016. doi:10.1136/annrheumdis-2016-209686
2 Hoffman HM, Wanderer AA. Inflammasome and IL-1beta-mediated disorders. *Curr Allergy Asthma Rep* 2010;10:229–35.
3 Aróstegui JI, Lopez Saldaña MD, Pascal M, *et al*. A somatic NLRP3 mutation as a cause of a sporadic case of chronic infantile neurologic, cutaneous, articular syndrome/neonatal-onset multisystem inflammatory disease: novel evidence of the role of low-level mosaicism as the pathophysiologic mechanism underlying Mendelian inherited diseases. *Arthritis Rheum* 2010;62:1158–66.

Response to: 'Criteria for CAPS, is it all in the name?' by Touitou and Sarrabay

We thank the authors of the eletter 'Criteria for CAPS, is it all in the name?'[1] for her very thoughtful comments on how to best define and name the autoinflammatory diseases we currently call cryopyrin-associated periodic syndromes—CAPS. The author very eloquently elaborate on some of the challenges we face when caring for children and adults with autoinflammatory diseases. Traditionally, disease names have highlighted the specific clinical phenotype—familiar cold-induced urticaria—or marked the name of the physicians, who first recognised a distinct disease entity—Muckle-Wells syndrome, Kawasaki disease. The discovery of a common genetic cause for the spectrum of illnesses we now label CAPS has changed this approach assuming a solid genotype–phenotype correlation. Subsequent disease names have followed this path—STING-associated vasculopathy with onset in infancy, haploinsufficiency of A20 and others. We may have hopefully assumed that patients with the clinical diagnosis of CAPS *always* have a disease-causing genetic variant in the *NLRP3* gene and vice versa. The author summarises carefully the diversity of genes associated with the diversity of clinical phenotypes in CAPS.

Our team chose a clinical approach. All team members care for children and adults with autoinflammatory diseases in different parts of the world and conduct research to improve the disease outcomes. The desire of the team was to enable a rapid diagnosis of the diseases in the clinical spectrum of CAPS to initiate treatment and prevent irreversible organ diseases. When conducting this rigorous exercise, we rejected the idea of classification criteria for research studies and the consideration of a mandatory genetic confirmation. For validation we did not choose genetically defined cohorts; the focus was the clinical phenotype and the true differential diagnosis in clinical practice.[2]

There are limitations to any approach we choose to take. For our team, children and adults suffering from the clinical phenotype of CAPS were the priority. We are all aware that confirming a diagnosis and giving a name is often the key for access to life and organ saving, expensive medications. The discussion around the best name for autoinflammatory diseases—capturing clinical entities versus genetic entities—is important. We thank the author again for her very thoughtful letter.

Jasmin B Kuemmerle-Deschner,[1] Susanne M Benseler[2]

[1]Division of Pediatric Rheumatology, Department of Pediatrics, University Hospital Tübingen, Tübingen, Germany
[2]Division of Rheumatology, Department of Pediatrics, Alberta Children's Hospital/University of Calgary, Calgary, Alberta, Canada

Correspondence to Dr Jasmin B Kuemmerle-Deschner, Division of Pediatric Rheumatology, Department of Pediatrics, University Hospital Tübingen, Hoppe-Seyler-Str. 3, Tübingen 72076, Germany; kuemmerle.deschner@uni-tuebingen.de

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

To cite Kuemmerle-Deschner JB, Benseler SM. *Ann Rheum Dis* 2017;**76**:e10.

Received 7 November 2016
Revised 8 November 2016
Accepted 9 November 2016
Published Online First 29 November 2016

► http://dx.doi.org/10.1136/annrheumdis-2016-210681

Ann Rheum Dis 2017;**76**:e10. doi:10.1136/annrheumdis-2016-210726

REFERENCES

1 Touitou I, Sarrabay G. Criteria for CAPS, is it all in the name? *Ann Rheum Dis* 2017;76:e9.
2 Kuemmerle-Deschner JB, Ozen S, Tyrrell PN, *et al*. Diagnostic criteria for cryopyrin-associated periodic syndrome (CAPS). *Ann Rheum Dis* 2016; Published Online First 4 October 2016.

PP2A plays a key role in inflammation and cancer through tristetraprolin activation

We have read with great interest the recent work by Ross *et al*,[1] which provides novel relevant findings about the therapeutic efficacy of using protein phosphatase 2A (PP2A)-activating drugs to target tristetraprolin (TTP) in rheumatoid arthritis (RA). In this elegant work, the authors showed that TTP is overexpressed and colocalises with activated mitogen-activated protein kinase (MAPK) p38 in RA synovial tissue. MAPK p38 phosphorylates and inhibits TTP at two serine residues, and Ross *et al* determined that these phosphorylation sites are critical for the role of TTP as a key regulator of inflammatory responses. Since PP2A dephosphorylates TTP at these two serine residues,[2] they hypothesised and assessed the efficacy of PP2A-activating compounds such as COG1410 and ALL(s) in RA, observing that these agents led to PP2A-mediated TTP activation thereby reducing both inflammation and bone erosion using in vitro and in vivo models of this disease. These results highlight the potential clinical usefulness of PP2A activation as a novel strategy to develop potent anti-inflammatory treatments.

PP2A is a well-known tumour suppressor that has been described commonly inactivated in human cancer.[3] Moreover, PP2A has been described as a key regulator of the MAPK signalling[4] and controls the production of proinflammatory chemokines.[5] Of importance, the risk of developing cancer is higher in people with inflammatory diseases such as colitis or hepatitis. The reason may be that the molecular events generated by the inflammatory response predispose to transformation from chronic inflammation to neoplasia.[6] Thus, TTP could represent a key linker between inflammation and cancer due to its role as modulator of the expression of both cytokines and proto-oncogenes.[7] Therefore, the function of PP2A as a TTP activator could be of high relevance and further reinforced by the fact that this phosphatase also targets MAPK signalling, which is responsible for the inhibitory phosphorylation of TTP.[5]

In conclusion, the study by Ross *et al*[1] highlights that PP2A plays a relevant role in inflammation through TTP dephosphorylation and activation. Importantly, these findings would suggest the potential benefits derived from the clinical use of PP2A-activating drugs as anti-inflammatory therapy as well as a novel strategy to prevent cancer development in those patients with chronic inflammatory diseases.

Ion Cristóbal,[1] Blanca Torrejón,[1] Juan Madoz-Gúrpide,[2] Federico Rojo,[2] Jesús García-Foncillas[1]

[1]Translational Oncology Division, Oncohealth Institute, IIS-Fundacion Jimenez Diaz, UAM, University Hospital 'Fundacion Jimenez Diaz', Madrid, Spain
[2]Pathology Department, University Hospital 'Fundacion Jimenez Diaz', Autonomous University of Madrid, Madrid, Spain

Correspondence to Dr Ion Cristobal and Professor Jesús García-Foncillas, Translational Oncology Division, Oncohealth Institute, IIS-Fundacion Jimenez Diaz, UAM, University Hospital "Fundacion Jimenez Diaz", Avda. Reyes Católicos-2, E-28040 Madrid, Spain; ion.cristobal@fjd.es, jgfoncillas@fjd.es

Contributors IC and BT wrote the manuscript. JM-G, FR and JG-F helped to draft the manuscript.

Funding This work was supported by PI13/02609 and PI15/00934 grants from 'Instituto de Salud Carlos III FEDER'. BT is supported by 'Fundación Conchita Rábago de Jiménez Díaz'.

Competing interests None.

Patient consent Obtained.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Cristóbal I, Torrejón B, Madoz-Gúrpide J, *et al*. *Ann Rheum Dis* 2017;**76**:e11.

Received 14 October 2016
Accepted 18 October 2016
Published Online First 3 November 2016

Ann Rheum Dis 2017;**76**:e11. doi:10.1136/annrheumdis-2016-210684

REFERENCES

1 Ross EA, Naylor AJ, O'Neil JD, *et al*. Treatment of inflammatory arthritis via targeting of tristetraprolin, a master regulator of pro-inflammatory gene expression. *Ann Rheum Dis* 2017;76:612–9.
2 Sun L, Stoecklin G, Van Way S, *et al*. Tristetraprolin (TTP)-14-3-3 complex formation protects TTP from dephosphorylation by protein phosphatase 2a and stabilizes tumor necrosis factor-alpha mRNA. *J Biol Chem* 2007;282:3766–77.
3 Sangodkar J, Farrington CC, McClinch K, *et al*. All roads lead to PP2A: exploiting the therapeutic potential of this phosphatase. *FEBS J* 2016;283:1004–24.
4 Westermarck J, Hahn WC. Multiple pathways regulated by the tumor suppressor PP2A in transformation. *Trends Mol Med* 2008;14:152–60.
5 Sandler H, Stoecklin G. Control of mRNA decay by phosphorylation of tristetraprolin. *Biochem Soc Trans* 2008;36:491–6.
6 Compare D, Nardone O, Nardone G. Non-steroidal anti-inflammatory drugs in the carcinogenesis of the gastrointestinal tract. *Pharmaceuticals (Basel)* 2010;3:2495–516.
7 Sanduja S, Blanco FF, Young LE, *et al*. The role of tristetraprolin in cancer and inflammation. *Front Biosci (Landmark Ed)* 2012;17:174–88.