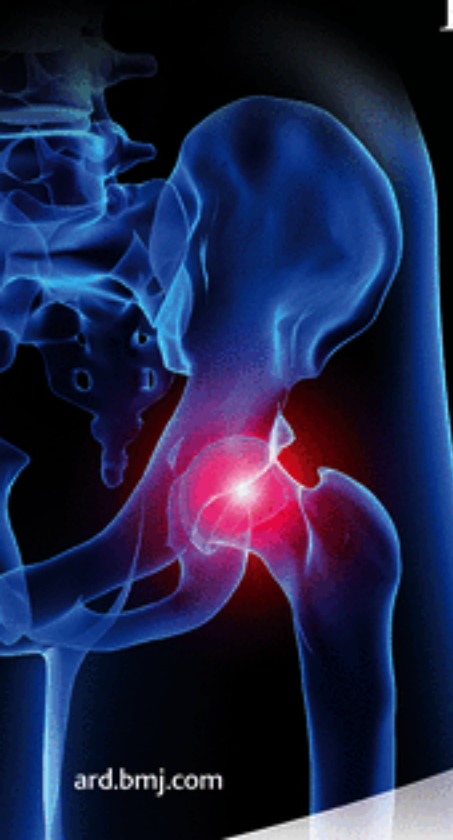
November 2017 Volume 76 Issue 11

Impact Factor 12.811

Annals of the Rheumatic Diseases

The EULAR Journal



ard.bmj.com

eular 70 BMJ

ARD
The EULAR Journal

Guidelines for Authors and Reviewers

Full instructions are available online at http://ard.bmj.com/pages/authors. Articles must be submitted electronically at http://mc.manuscriptcentral.com/ard. Authors retain copyright but are required to grant ARD an exclusive licence to publish. (http://authors.bmj.com/submitting-your-paper/copyright-and-authors-rights).

Annals of the Rheumatic Diseases publishes original work on all aspects of rheumatology and disorders of connective tissue. Laboratory and clinical studies are equally welcome

Contact Details

Editorial Office
Annals of the Rheumatic Diseases
BMJ Publishing Group Ltd
BMA House
Tavistock Square
London WCIH 9JR,UK
T: **+44 (0)20 7383 6250**
E: **ard@bmj.com**

Permissions
http://www.bmj.com/company/products-services/rights-and-licensing/permissions

Supplement Enquiries
T: **+44 (0)20 7383 6057**
E: **journals@bmj.com**

Subscriptions
For all subscription enquiries and orders
T: **+44 (0)20 7111 1105**
E: **http://ard.bmj.com/pages/subscribe**

Display Advertising Sales
Sophie Fitzsimmons
T: **+44 (0)20 7383 6783**
E: **sfi tzsimmons@bmj.com**
http://www.bmj.com/company/raise-visibility-and-reach

Online Advertising Sales
Marc Clifford
T: **+44 (0) 20 7383 6161**
E: **mclifford@bmj.com**
http://www.bmj.com/company/raise-visibility-and-reach

Display & Online Advertising Sales (USA)
American Medical Communications (AMC)
T: **+1 732 490 5530**
E: **jloughran@americanmedicalcomm.com**

Author Reprints
Reprints Administrator
E: **admin.reprints@bmj.com**

Commercial Reprints (except USA & Canada)
Nadia Gurney-Randall
T: **+44 (0)20 8445 5825**
M: **+44 07866 262344**
E: **ngurneyrandall@bmj.com**

Commercial Reprints (USA & Canada)
Ray Thibodeau
T: **+1 267 895 1758**
M: **+1 215 933 8484**
E: **ray.thibodeau@contentednet.com**

EULAR
Eular Executive Secretariat
Seestrasse 240, 8802 Kilchberg, Switzerland
E: **eular@eular.org**
www.eular.org

For all other ARD contacts
http://ard.bmj.com/contact-us

Subscription Information

ARD is published monthly; subscribers receive all supplements
ISSN 0003-4967 (print); 1468-2060 (online)

Institutional Rates 2017

Print
£939

Online
Site licences are priced on FTE basis and allow access by the whole institution. Details available online at http://journals.bmj.com/content/subscribers or contact the Subscription Manager in the UK (see above right)

Personal Rates 2017

Print (includes online access at no additional cost)
£387

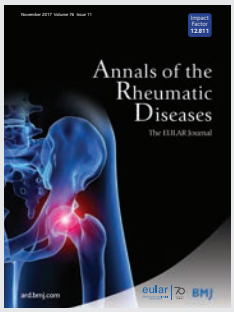
Online only
£164

Eular congress delegates
Delegates receive a Continuous Professional Development package that includes a 12 month complimentary subscription to *ARD* in print and/or online

Personal print or online only and institutional print subscriptions may be purchased online at http://journals.bmj.com/content/subscribers (payment by Visa/Mastercard only)

Residents of some EC countries must pay VAT; for details, call us or visit http://journals.bmj.com/content/subscribers

ARD
The Eular Journal



Editorial office
Annals of the Rheumatic Diseases
BMJ Publishing Group Ltd
BMA House
Tavistock Square
London WCIH 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: 12.811

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

Recommendation

Clinical and epidemiological research

MORE CONTENTS ▶

Basic and translational research

Electronic pages

Correspondence

Glucocorticoids in rheumatoid arthritis: the picture is shaping up

Frank Buttgereit,[1] Johannes W Bijlsma[2]

Despite the well-known efficacy of glucocorticoids (GCs), it is more common than exceptional that their known adverse event profile and co-morbidity implications elicit fierce debates when discussing their benefit–risk profile. These discussions usually come up, especially among GC 'supporters' and 'opponents',[1] when trying to elaborate recommendations on how to use these drugs best in the treatment of rheumatic diseases such as rheumatoid arthritis (RA),[2] polymyalgia rheumatica,[3] giant cell arteritis,[4] systemic lupus erythematodes,[5 6] myositis[7] and even systemic sclerosis.[7] For example, very divergent opinions were learnt during work on the European League Against Rheumatism (EULAR) recommendations (2013 update) for the management of RA.[2] As a result, only 73% (the lowest majority level of all recommendations) of the members approved the suggestion that GCs should only be used as bridging therapy for up to 6 months, ideally tapering them at earlier time points. Interestingly, when looking at the members voting against it, about half of them thought the statement was too weak, whereas the other half considered it to be too strong.[2] Nevertheless, the level of agreement (strength of recommendation) was quite high (mean of 8.9) on final anonymous grading. It should also be noted that the group did not explicitly discuss the chronic use of GC in established RA.

Indeed, many patients and quite some physicians are still uncertain about the actual benefit:risk ratio of GCs.[1] This uncertainty might prevent optimal treatment under conditions where GC treatment is known to be of added value. Fortunately, our view on the most optimal use of these drugs does slowly but constantly maturate. This primarily results from thorough analyses of accumulated data in order to update recommendations, but there are also new and often qualitatively better data coming in to enrich our knowledge. A very good example of the latter is the work by Roubille *et al*.[8] These authors report carefully collected 7-year data from a prospective multicentre observational cohort of patients with early arthritis (ESPOIR). Current gaps in knowledge were addressed by analysing the tolerability profile of GC use in early arthritis. As a result, a comprehensive data set now lies on the table to be judged by everyone.

There are two key findings. First, the analysis of data of 602 patients with RA demonstrates patients with (versus without) GC treatment to be those with the greater use of non-steroidal anti-inflammatory drugs and disease modifying antirheumatic drugs, more active disease and higher C-reactive protein and anti-citrullinated protein antibody levels. These data confirm that there may be a significant bias in form of confounding by indication when analysing GC effects. Roubille and colleagues properly addressed this problem by performing a weighted Cox proportional-hazards analysis, with the use of propensity score and inverse probability-of-treatment weighting, and including age, gender, history of hypertension and GC treatment. When doing so, the real-life tolerability outcomes did not show any significant difference between RA patients with and without GC treatment. The facts that most of the patients who took GCs started this therapy during the first 6 months and that they received <5 mg prednisone per day clearly support the good safety profile of low-dose GCs for early active RA. Second, the mean dosage during the entire follow-up was only 3.1±2.9 mg/day which represents a very low-dose GC treatment.[9] Therefore, it should be stressed that these safety data do not automatically lend support to higher doses.

HOW CAN WE ACHIEVE A POSITIVE BENEFIT–RISK PROFILE WHEN USING GLUCOCORTICOIDS IN THE TREATMENT OF PATIENTS WITH RHEUMATOID ARTHRITIS?

The real-world data obtained by the French group fit perfectly with results recently published. A multidisciplinary EULAR group of experts including patients with rheumatic diseases aimed to define conditions under which long-term (3–6 months or longer) GC therapy has an acceptably low level of harm.[2] Following a thorough analysis for the four most worrisome GC adverse effects (on bone, hyperglycaemia/diabetes mellitus, cardiovascular diseases and infections), the group concluded that the risks of long-term GC therapy are defined by drug-specific parameters (dose and duration) but at least as much by patient-specific characteristics (eg, age, gender, genetic pre-disposition, co-morbidities, co-medication and individual lifestyle) (figure 1). Although robust evidence on the risk of harm was often lacking, long-term dosages of ≤5 mg prednisone equivalent per day can be considered to have an acceptably low level of harm. This appears to be true for the majority of patients with the exception of patients at high risk of cardiovascular disease who may require preventive measures.[2]

THE UPDATED VIEW ON USING GLUCOCORTICOIDS IN THE TREATMENT OF RHEUMATOID ARTHRITIS

The view of the 2013 update of RA recommendations was that both short-term symptomatic and long-term structural effects of GCs should be used in the form of a bridging therapy. Therefore, these drugs should be administered as long as it takes for conventional synthetic disease-modifying antirheumatic drugs (DMARDs) to reach their maximum effect, but should then be tapered if clinically feasible. This view has been further evolved in the most recent update of these recommendations.[10] The new wording—a compromise in order to accommodate most of the concerns and suggestions raised during the Task Force's debate—is as follows (recommendation 6): 'Short-term glucocorticoids should be considered when initiating or changing conventional synthetic DMARDs, in different dose regimens and routes of administration, but should be tapered as rapidly as clinically feasible.' This wording does acknowledge existing differences in GC application in terms of dose regimen and routes, that is, there are several different regimens for oral use, intramuscular injection and intravenous pulse therapy. It is now also being stated more clearly that GCs should be given as bridging therapy together with conventional synthetic DMARDs, either as part of the initial strategy or subsequently if this has failed. In contrast, GCs are usually not needed as a bridging

[1]Department of Rheumatology and Clinical Immunology, Charité—University Medicine Berlin, Berlin, Germany
[2]Department of Rheumatology & Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands

Correspondence to Professor Frank Buttgereit, Department of Rheumatology and Clinical Immunology, Charité – University Medicine Berlin, 10117 Berlin, Germany; frank.buttgereit@charite.de

therapy when biological or targeted synthetic DMARDs are used. The reason is that these drugs typically have a rapid onset of action and the infection risks may be potentiated.[11 12]

It should be noted that the second part of this recommendation has been left unchanged stressing again that GCs should be gradually reduced and ultimately stopped, ideally within 3–6 months.[10] The main reason for this is that the long-term use of GCs, especially at doses >5 mg/day, should be avoided because of the potential occurrence of adverse effects.[1] In addition, the Task Force explicitly stated that the DMARD therapy may have to be considered a failure (and is, therefore, in need of optimisation) if GCs cannot be withdrawn within the time frame mentioned above.

In our opinion, chronic use of low-dose GCs can be considered a realistic option for some patients, based on the following observations:

1. Recent data show that a large proportion of patients with early RA do stay on GC therapy for >6 months, indicating an acceptable balance between efficacy and safety. For example, the data reported by Roubille et al[8] show that 64% of RA patients received low-dose prednisone for the entire follow-up. The mean duration of total GC treatment was 1057±876 days, which is much longer than the recommended maximum period of 6 months. Data recently published on the routine care of early RA point were in the same direction.[13] More than 1300 patients were followed for up to 2 years in an early arthritis cohort (CAPEA), and complete 2-year data were available for 669 patients with RA. Seventy-seven per cent of those patients were initially treated with GC at different starting doses (26% <7.5 mg, 29% 7.5–20 mg and 45% ≥20 mg of prednisolone per day). Of note, after 24 months, 47% still were on GC.
2. Also in established RA, there are many patients obviously being more or less constantly treated with GC. For example, recently published German data from the National Database of the Collaborative Arthritis describe a total of 8084 patients with RA. Forty-eight per cent of these patients received a mean dose of 5 mg prednisone equivalent per day, whereby 8.5% were treated with daily dosages <5 mg, 37.7% with 5–7.5 mg and 2.1% with >7.5 mg.[14]
3. As outlined above, for many patients, a therapy with dosages ≤5 mg prednisone equivalent for >6 months may have an acceptably low level of harm and will provide therapeutic effects on top of the DMARD therapy. This assumption is backed by both the published EULAR Task Force work[1] and the real-life data by Roubille et al.[8]

In conclusion, during treatment of both early and established RA, the risks of adverse effects induced by conventional GC can be minimised when following the established recommendations,[15–17] and by considering each patient to be an individual person characterised by the presence or absence of certain risk factors and/or preventive measures. This will ultimately result in an adapted patient-specific therapy. Drawing updated conclusions based both on new data coming in and on balanced analyses of data already existing is a better way to go than relying on eminence-based statements (as has often been observed in the past and provided by both GC 'supporters' and 'opponents'). Not picking single observations or speculating on limited (and/or sometimes biased) data sets in order to convey certain negative or positive judgements on these drugs, but rather full and continuous analyses of the whole picture are what is needed. Glucocorticoids are (still) just too important to make half-hearted, opinion-driven statements.

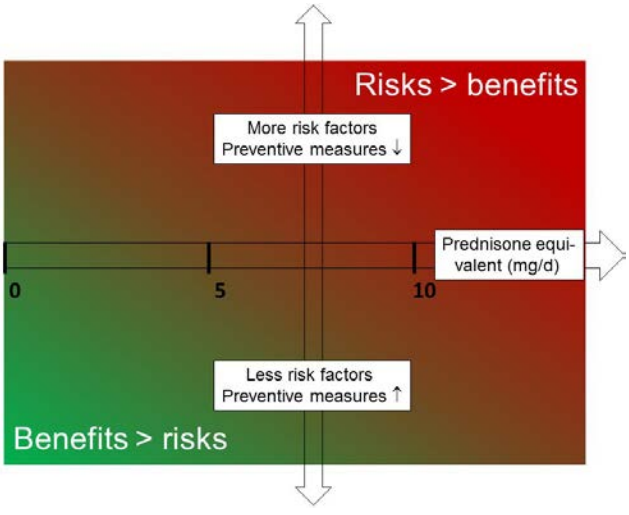


Figure 1 A matrix illustrating the benefit:risk ratio for a long-term therapy with GC. In general, lower GC doses together with the presence of less risk factors and more preventive measures lead to a lower level of harm, thereby leading to a better benefit:risk ratio of such a treatment (the green area). In contrast, higher GC doses together with the presence of more risk factors and less preventive measures lead to a higher level of harm, thereby leading to a worse benefit:risk ratio (the red area). More specifically, a therapy with dosages ≤5 mg prednisone equivalent for >6 months may have an acceptably low level of harm in the majority of patients (with exception of patients at high risk of CVD who may require preventive measures). At dosages between >5 and ≤10 mg/day, an acceptably low level of harm can only be assumed in the absence of certain risk factors and/or if appropriate preventive measures are taken. That is why the green colour gradually turns into red from bottom left to upper right. GC, glucocorticoids.

Contributors Both authors have written together this article.

Competing interests FB has received consultancy fees, honoraria and travel expenses from Horizon Pharma, Mundipharma Int. Ltd. and Roche, and grant support from Mundipharma Int. Ltd. and Horizon Pharma. JWJB has received consultancy fees from Horizon, Mundipharma, Enceladus and SUN.

Provenance and peer review Commissioned; externally peer reviewed.

To cite Buttgereit F, Bijlsma JW. *Ann Rheum Dis* 2017;**76**:1785–1787.

► http://dx.doi.org/10.1136/annrheumdis-2016-210135

Ann Rheum Dis 2017;**76**:1785–1787. doi:10.1136/annrheumdis-2017-211187

REFERENCES

1 Strehl C, Bijlsma JW, de Wit M, *et al*. Defining conditions where long-term glucocorticoid treatment has an acceptably low level of harm to facilitate implementation of existing recommendations: viewpoints from an EULAR task force. *Ann Rheum Dis* 2016;75:952–7.

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 Dejaco C, Singh YP, Perel P, *et al*. 2015 recommendations for the management of polymyalgia rheumatica: a European league against rheumatism/American college of rheumatology collaborative initiative. *Ann Rheum Dis* 2015;74:1799–807.
4 Hoff P, Gaber T, Strehl C, *et al*. Immunological characterization of the early human fracture hematoma. *Immunol Res* 2016;64:1195–206.
5 Aringer M, Leuchten N, Fischer-Betz R. [Tapering and termination of immunosuppressive therapy : systemic lupus erythematosus]. *Z Rheumatol* 2017;76:27–32.
6 Apostolopoulos D, Morand EF. It hasn't gone away: the problem of glucocorticoid use in lupus remains. *Rheumatology* 2016:kew406.
7 Postolova A, Chen JK, Chung L. Corticosteroids in myositis and scleroderma. *Rheum Dis Clin North Am* 2016;42:103–18.
8 Roubille C, Rincheval N, Dougados M, *et al*. Seven-year tolerability profile of glucocorticoids use in early rheumatoid arthritis: data from the ESPOIR cohort. *Ann Rheum Dis* 2017;76:1797–802.
9 Buttgereit F, da Silva JA, Boers M, *et al*. Standardised nomenclature for glucocorticoid dosages and glucocorticoid treatment regimens: current questions and tentative answers in rheumatology. *Ann Rheum Dis* 2002;61:718–22.
10 Smolen JS, Landewé R, Bijlsma J, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76:960–77.
11 Listing J, Kekow J, Manger B, *et al*. Mortality in rheumatoid arthritis: the impact of disease activity, treatment with glucocorticoids, tnfα inhibitors and rituximab. *Ann Rheum Dis* 2015;74:415–21.
12 Lahiri M, Dixon WG. Risk of infection with biologic antirheumatic therapies in patients with rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2015;29:290–305.
13 Albrecht K, Callhoff J, Edelmann E, *et al*. [Clinical remission in rheumatoid arthritis. data from the early arthritis cohort study CAPEA]. *Z Rheumatol* 2016;75:90–6.
14 Albrecht K, Huscher D, Eidner T, *et al*. [Medical treatment of rheumatoid arthritis in 2014 : current data from the german collaborative arthritis centers]. *Z Rheumatol* 2017;76:50–7.
15 Palmowski Y, Buttgereit T, Dejaco C, *et al*. The "official view" on glucocorticoids in rheumatoid arthritis. A systematic review of international guidelines and consensus statements. *Arthritis Care Res* 2016.
16 van der Goes MC, Jacobs JW, Boers M, *et al*. Monitoring adverse events of low-dose glucocorticoid therapy: eular recommendations for clinical trials and daily practice. *Ann Rheum Dis* 2010;69:1913–9.
17 Da Silva JA, Jacobs JW, Kirwan JR, *et al*. Safety of low dose glucocorticoid treatment in rheumatoid arthritis: published evidence and prospective trial data. *Ann Rheum Dis* 2006;65:285–93.

European evidence-based recommendations for diagnosis and treatment of childhood-onset systemic lupus erythematosus: the SHARE initiative

Noortje Groot,[1,2] Nienke de Graeff,[1] Tadej Avcin,[3] Brigitte Bader-Meunier,[4] Paul Brogan,[5] Pavla Dolezalova,[6] Brian Feldman,[7] Isabelle Kone-Paut,[8] Pekka Lahdenne,[9] Stephen D Marks,[5] Liza McCann,[10] Seza Ozen,[11] Clarissa Pilkington,[5] Angelo Ravelli,[12] Annet van Royen-Kerkhof,[1] Yosef Uziel,[13] Bas Vastert,[1] Nico Wulffraat,[1] Sylvia Kamphuis,[2] Michael W Beresford[10,14]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210960).

For numbered affiliations see end of article.

Correspondence to
Dr Noortje Groot, Department of Paediatric Immunology, University Medical Centre Utrecht, Lundlaan 6, Utrecht 3584 EA, The Netherlands; n.groot@erasmusmc.nl

NG and NG contributed equally, SK and MWB contributed equally.

Received 12 December 2016
Revised 2 May 2017
Accepted 4 May 2017
Published Online First 19 June 2017

ABSTRACT

Childhood-onset systemic lupus erythematosus (cSLE) is a rare, multisystem and potentially life-threatening autoimmune disorder with significant associated morbidity. Evidence-based guidelines are sparse and management is often based on clinical expertise. SHARE (Single Hub and Access point for paediatric Rheumatology in Europe) was launched to optimise and disseminate management regimens for children and young adults with rheumatic diseases like cSLE. Here, we provide evidence-based recommendations for diagnosis and treatment of cSLE. In view of extent and complexity of cSLE and its various manifestations, recommendations for lupus nephritis and antiphospholipid syndrome will be published separately. Recommendations were generated using the EULAR (European League Against Rheumatism) standard operating procedure. An expert committee consisting of paediatric rheumatologists and representation of paediatric nephrology from across Europe discussed evidence-based recommendations during two consensus meetings. Recommendations were accepted if >80% agreement was reached. A total of 25 recommendations regarding key approaches to diagnosis and treatment of cSLE were made. The recommendations include 11 on diagnosis, 9 on disease monitoring and 5 on general treatment. Topics included: appropriate use of SLE classification criteria, disease activity and damage indices; adequate assessment of autoantibody profiles; secondary macrophage activation syndrome; use of hydroxychloroquine and corticosteroid-sparing regimens; and the importance of addressing poor adherence. Ten recommendations were accepted regarding general diagnostic strategies and treatment indications of neuropsychiatric cSLE. The SHARE recommendations for cSLE and neuropsychiatric manifestations of cSLE have been formulated by an evidence-based consensus process to support uniform, high-quality standards of care for children with cSLE.

INTRODUCTION

Childhood-onset systemic lupus erythematosus (cSLE) is a severe, chronic, systemic autoimmune disease that has great impact on the child or young person affected. cSLE shares its pathogenesis with adult-onset SLE, but generally has a more severe clinical phenotype.[1–8]

With an incidence of 0.3–0.9 per 100 000 children-years and a prevalence ranging from 1.89 to 25.7 per 100 000 children worldwide (reviewed in refs[9–11]), including Europe,[12–16] cSLE fulfils the definition of a rare disease in Europe.[17] Its low prevalence makes clinical research challenging, resulting in a paucity of evidence-based data and subsequent guidelines for disease management. Consequently, the management of patients with cSLE differs widely between countries.[18] Treatment approaches can vary between clinicians even within centres. To foster equity of access to the highest standards of care and uniformity of practice, evidence-based international guidelines are therefore urgently needed.

To achieve this, collaboration between countries is necessary. For this reason, the SHARE (Single Hub and Access point for paediatric Rheumatology in Europe) project was initiated.[18] One of the key objectives of this project was to provide guidance regarding best practices for the diagnosis and management of paediatric rheumatic diseases. SHARE recommendations for autoinflammatory diseases and juvenile dermatomyositis have been published.[19–21] Here, we present SHARE recommendations for cSLE. In view of extent and complexity of cSLE, SHARE recommendations for lupus nephritis (LN) and antiphospholipid syndrome (APS) will be published separately.

METHODS

A European-wide panel of 16 paediatric rheumatologists and representation of paediatric nephrology was established to develop evidence-based recommendations. A project plan for the systematic literature search was written following the EULAR (European League Against Rheumatism) standardised operating procedure.[22] SHARE was a European Union (EU)-funded project and as such there was a prerequisite for representative disease experts from across Europe to form the expert panel, with inclusion of a selected number of disease experts from outside the EU.

Systematic literature search and study selection

A systematic literature search based on specific research questions was performed in PubMed/MEDLINE, EMBASE and Cochrane databases

To cite: Groot N, de Graeff N, Avcin T, *et al*. *Ann Rheum Dis* 2017;**76**:1788–1796.

in July 2013 (see online supplementary table S1). A validated filter was used to specifically select articles on children and adolescents only.[23] The filter was adapted for cSLE to exclude neonates, as neonatal lupus was beyond the scope of the review (see online supplementary table S2). The literature search also included LN and paediatric APS. These topics will be discussed separately from this article.

Validity assessment

Two reviewers (NG, NdG) independently screened all articles according to the predefined inclusion and exclusion criteria.

Articles were reviewed independently by two cSLE experts from European panel (MWB, SK, TA, AR, IKP, BBM, CP). They assessed level of evidence and methodological quality of the articles (see online supplementary tables S3 and S4).[24 25] Data extraction was done by the experts using predefined data extraction forms adapted from classification tables for epidemiologic, diagnostic[26] and therapeutic[27] studies. If there were any discrepancies, a third expert was asked to give a final assessment.

Establishment of recommendations

The results of the literature review were mapped against the a priori research questions, and provisional recommendations were formulated (NG, NdG, SK, MWB). If no literature in children could be found to map against a particular recommendation, adult literature was consulted. The expert committee (TA, BBM, PB, PD, IKP, PL, LM, SO, CP, AR, AvR, YU, NW, SK, MWB) was presented with the provisional recommendations in web-based surveys (100% response rate) and gave their opinions on the statements. Recommendations were revised according to responses to the surveys and discussed at two sequential face-to-face consensus meetings in March 2014 (Genova, number of experts participating, n=15; moderators: BF and AR) and March 2015 (Barcelona, n=14; moderator: BF).

To reach consensus, the nominal group technique was used, in which equal participation from group members is ensured.[28] Recommendations were accepted when agreement was at least 80%. This process resulted in a final set of prioritised recommendations.

RESULTS

Figure 1 summarises the results of the literature search. A total of 9341 articles were identified and reviewed regarding treatment and management of cSLE, of which 133 articles fulfilled the inclusion criteria. We identified 51 articles relating to diagnosis and management of cSLE generally, and 27 articles to neuropsychiatric manifestations of cSLE, all were scored by the experts (see online supplementary table S1). The 55 articles pertaining to LN informed a specific set of complimentary recommendations that will be published separately.

The meetings resulted in 25 recommendations pertaining to the diagnosis and treatment of cSLE (table 1) and 10 for neuropsychiatric cSLE (NP-cSLE) (table 2). The recommendations in this paper can be used for all patients in whom cSLE is suspected or diagnosed.

The most severe symptom(s) or sign(s) should guide treatment decisions when considering these recommendations. For example, when a patient suffers from mild haematological involvement as well as severe neuropsychiatric disease, the latter should guide the treatment choice.

General diagnostic recommendations

Prompt, accurate diagnosis of cSLE in a specialist centre is crucial to enable timely initiation of appropriate treatment, including multidisciplinary care. However, there are no validated diagnostic criteria for cSLE. Despite some differences regarding symptoms at onset, pattern of organ involvement and severity of disease between cSLE and adult-onset disease,[3 29] their similarities mean that the established American College of Rheumatology (ACR) classification criteria for SLE are widely used for cSLE.[30] In 2012, new classification criteria for SLE were published.[31] To date, two studies have assessed the performance of these Systemic Lupus International Collaborating Clinics (SLICC) classification criteria for SLE in children.[6 7] Both concluded that although some specificity may be lost, the SLICC criteria had better sensitivity than the ACR criteria. Evidence to date indicates the SLICC criteria may well be preferable in cSLE, and should be used to aid referral to, or at least consultation with a paediatric rheumatologist. Similarly, they may help prompt referral, even if a child does not yet meet full criteria, since these are classification and not diagnostic criteria.

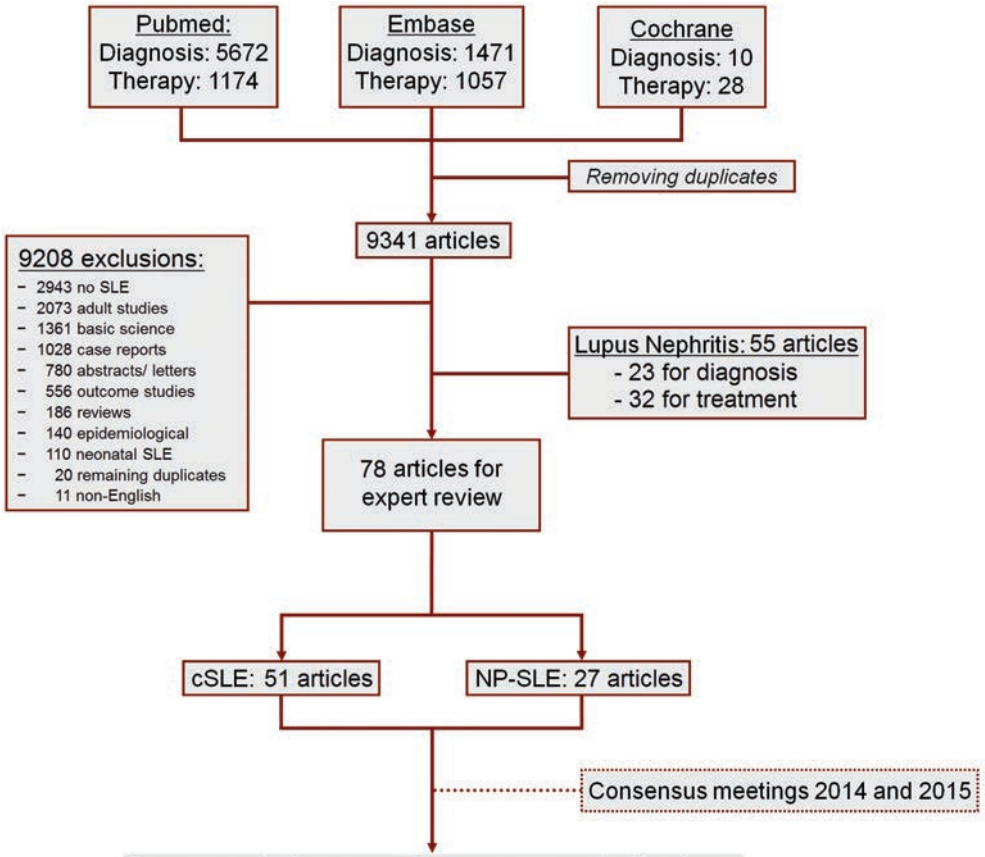
A hallmark of SLE is the presence of autoantibodies, particularly those directed towards nuclear autoantigens (antinuclear antibodies, ANA). Next to ANA, autoantibodies including anti-double-stranded DNA (anti-dsDNA), anti-Sm, anti-RNP, anti-Ro/SS-A and anti-La/SS-B (collectively referred to as 'ENA' (extractable nuclear antigens)) are prevalent in cSLE: dsDNA 54%–93%; anti-Sm 17%–52%; anti-RNP 22%–50%; anti-Ro/SS-A 33%–54%; anti-La/SS-B 14%–32%.[5 32–38] As such, including all of these antibodies in the diagnostic work-up when considering cSLE was strongly recommended. However, there are no antibodies with specific predictive qualities (eg, disease severity, organ involvement, age of onset) despite extensive efforts to find them.[39–47] Notably, patients negative for anti-dsDNA antibodies and/or ENA can still be diagnosed with cSLE.

Hereditary complement deficiencies can predispose to lupus or lupus-like disease at an early age. Early recognition of these deficiencies should facilitate adequate treatment of disease and comorbidities including infections, which are especially important as these patients seem to have a higher mortality.[48 49] Therefore, screening for complement deficiencies via CH50, AP50, C3 and C4 (or other validated classic and alternative complement pathway assay) is important in cSLE, especially in young patients with lupus. It was also recognised that there are other causes of monogenic lupus outside of the complement pathway, thus normal complement screening assay results do not preclude this possibility.[50 51]

Cardiopulmonary involvement

Although unusual in cSLE, cardiac and pulmonary involvement does occur, but is often asymptomatic initially.[52–60] Respiratory symptoms or signs such as exertional intolerance could be a sign of pulmonary or cardiac pathology. However, there is a wide differential diagnosis that must be considered and use of appropriate diagnostic procedures should consequently be performed to find out whether cardiopulmonary involvement is due to cSLE.

Early recognition of cardiopulmonary involvement is important when trying to prevent subsequent organ damage. Therefore, a baseline echocardiography and ECG screen in every patient with cSLE for cardiopulmonary involvement is advised. Additionally, intermittent monitoring for any future progression or new involvement of these organ systems over time can be considered, as it is not clear how many children with asymptomatic cardiopulmonary involvement become symptomatic.



Recommendations regarding:	Number
cSLE	
Diagnosis	11
Management	9
Treatment	5
NP-cSLE	
Diagnosis	6
Management	4

Figure 1 Summary of the literature search. cSLE, childhood-onset systemic lupus erythematosus; LN, lupus nephritis; NP-cSLE, neuropsychiatric cSLE; NP-SLE, neuropsychiatric SLE; SLE, systemic lupus erythematosus.

Macrophage activation syndrome

Macrophage activation syndrome (MAS) is a rare but severe, potentially life-threatening complication of cSLE, characterised by high fever, associated in some patients with organ involvement (neurological symptoms, hepatomegaly, splenomegaly), pancytopenia, coagulopathy, elevation of liver enzymes, ferritin and triglycerides.[61 62] Preliminary recommendations for timely diagnosis and correct classification of MAS in cSLE have been developed.[61] Patients can develop MAS at any time during their disease. Distinguishing sepsis from MAS can be difficult, as they may share features such as fever, cytopenias and hepatic involvement, both resulting in systemic inflammation. There are differences, as for example ferritin levels in MAS tend to increase dramatically, whereas hyperferritinaemia is generally more modest in sepsis.[63 64] A bone marrow aspirate should be performed to assess the cause of cytopenias and to detect possible haemophagocytosis. This will help in making a diagnosis of MAS. As MAS can be rapidly progressive and life threatening, the threshold for diagnostic procedures should be low. However, if the patient is clinically unstable, treatment should not be delayed if a bone marrow aspirate is not possible.

Monitoring and general management

The frequency of visits to the paediatric outpatient clinic is dependent on clinical presentation, disease severity as well as the age of the patient. Visits should be regular, especially at diagnosis and following flares and a basic set of investigations is recommended for each visit.[65 66] Consensus was reached on preliminary criteria for global flares in cSLE.[67] Further validation studies are needed to confirm the usefulness of the cSLE flare criteria in research and for clinical care. The recommended frequency of visits as well as the important clinical parameters that should be checked at each visit is similar to the recommendations for adult-onset disease.[68–70] In addition, regular height and weight monitoring is important, as well as pubertal assessment. Growth impairment can occur in children with cSLE, which can be difficult to overcome and may lead to a lower final height due to continuous disease activity and/or corticosteroid use. Similarly, these factors can contribute

Table 1 Recommendations for cSLE—diagnostic procedures, management and treatment

	Level of evidence	Strength	Agreement (%)
Diagnostic recommendations			
1. Based on the current evidence (mainly in adults) on the SLICC criteria, the SLICC criteria can be used as classification criteria in cSLE.	3	C	100
2. In the presence of a positive ANA combined with at least two clinical SLICC criteria, or in the presence of a positive ANA combined with at least one clinical and one immunological SLICC criterion, referral to a paediatric rheumatologist is warranted.	3	C	100
3. When considering a diagnosis of cSLE, anti-Sm, anti-RNP-a, anti-Ro/SS-A and anti-La/SS-B should be included routinely.	3	C	100
4. In a clinical context, when a patient is ANA positive, but anti-dsDNA and ENA negative, a diagnosis of cSLE can still be made.	3	C	100
5. In patients with cSLE, hereditary complement deficiencies should be considered, especially in young patients.	3	C	100
6. All patients with cSLE should have a chest X-ray at diagnosis.	3	C	100
7. All patients with cSLE should be screened for cardiac abnormalities using ECG and echocardiography at diagnosis.	3	C	100
8. Patients with cSLE with respiratory symptoms or signs (in the absence of acute infection) should have a pulmonary function test including CO diffusion.	3	C	100
9. Exertional intolerance in patients with cSLE should be investigated. Initial investigations should include a chest X-ray, a pulmonary function test (with CO diffusion), echocardiography and an ECG.	3	C	100
10. In patients with cSLE, unexplained fever should trigger a search for infection and MAS.	3	C	93
11. When MAS is suspected, a bone marrow aspirate should be considered to facilitate MAS diagnosis and exclude other diagnoses. If MAS is suspected and the patient is clinically unstable, treatment should not be delayed if a bone marrow aspirate is not possible.	3	C	100
Monitoring and management of cSLE			
1. Active disease should be regularly monitored by performing: a full clinical evaluation including body weight, height and blood pressure; urine dipstick testing; proteinuria estimation; blood tests including albumin; creatinine; eGFR; ESR; C3 and C4; anti-dsDNA; and complete blood cell count.*	2AB/3	B-C/ C	100
2. Clinical evaluation should usually occur every 2–4 weeks for the first 2–4 months after diagnosis or flare, and then according to the response to treatment.*	3	C	100
3. Children receiving systemic corticosteroids should be checked regularly for linear growth.	2A	B	100
4. All children with cSLE should have disease activity assessed regularly using a standardised disease activity measure such as the SLEDAI-2k or pBILAG-2004.	4	D	100
5. All children with cSLE should have disease damage assessed yearly using a standardised disease damage measure such as the paediatric SDI.	3	C	100
6. All patients with cSLE should have access to an ophthalmologist.	3	C	100
7. Annual eye screening should be considered for patients with cSLE taking hydroxychloroquine.	3	C	100
8. Sun protection may be beneficial in patients with skin manifestations and should be considered.*	3	C	100
9. In lupus, a coordinated transition programme including paediatric and adult specialists is crucial for ensuring continuity of care and adherence to treatments in order to optimise long-term outcome including prevention of fatalities.*	3	C	80
Treatment recommendations			
1. All children with lupus should be on hydroxychloroquine routinely.	2A	B	100
2. In all decisions of treatment change or modification, compliance should be actively checked.	3	C	100
3. When it is not possible to taper the prednisone dose, a DMARD should be added to the therapy.	3	C	100
4. Mild/moderate haematological involvement: when haemolysis is present and Hb is lower than normal, a DMARD should be added to the therapy.	3	C	100
5. If rituximab is required, the recommended dose is either 750 mg/m^2/dose (up to a maximum of 1 g) at day 1 and day 15, or 375 mg/m^2/dose once a week for four doses.	3	C	100

*This statement is based on the EULAR recommendations for adults with SLE.[68–70]
Level of evidence: for diagnostic and observational studies: 1A, meta-analysis of cohort studies;1B, randomised controlled study; 2A, controlled study without randomisation; 2B, quasiexperimental study; and for treatment studies: 1A, meta-analysis of randomised controlled trial; B, based on level 2 or extrapolated from level 1; C, based on level 3 or extrapolated from level 1 or 2; D, based on level 4 or extrapolated from level 3 or 4 expert opinion.[22]; Strength of recommendation: A, based on level 1 evidence; 3, descriptive study; 4, expert opinion.[25–27] Agreement indicates per cent of experts agreeing on the recommendation during the final voting round of the consensus meeting.
ANA, antinuclear antibodies; anti-dsDNA, anti-double-stranded DNA; CO, carbon monoxide; cSLE, childhood-onset systemic lupus erythematosus; DMARD, disease-modifying antirheumatic drug; eGFR, estimated glomerular filtration rate; ENA, extractable nuclear antigens; ESR, erythrocyte sedimentation rate; EULAR, European League Against Rheumatism; Hb, haemoglobin; MAS, macrophage activation syndrome; pBILAG-2004, paediatric British Isles Lupus Assessment Group index 2004; SDI, SLICC/American College of Rheumatology Damage Index; SLEDAI-2k, Systemic Lupus Erythematosus Disease Activity Index 2000; SLICC, Systemic Lupus International Collaborating Clinics.

to delayed pubertal development. Prepubertal and peripubertal patients receiving a high cumulative dose of corticosteroids are specifically at risk for both growth impairment and pubertal delay, which must be proactively assessed.[71 72]

It is strongly recommended that disease activity, response to treatment and disease damage should be regularly and comprehensively assessed using standardised tools to monitor disease progression. Many tools are available for this purpose.[73–75] For example, disease activity can be monitored with the paediatric

Table 2 Recommendations for NP-cSLE—diagnostic procedures and treatment

	Level of evidence	Strength	Agreement (%)
Diagnostic recommendations			
1. The nomenclature and case definitions for NP-cSLE syndromes proposed by the ACR ad hoc committee should be used to classify and describe NP-SLE syndromes in cSLE.	3	C	100
2. In patients with cSLE with new or unexplained symptoms or signs suggestive of neuropsychiatric disease, initial diagnostic work-up should include work-up as performed in patients without SLE.	3	C	100
3. In a patient with a suspected diagnosis of NP-cSLE or worsening NP-cSLE symptoms, underlying factors including infections, hypertension, metabolic abnormalities or adverse effects of medication should be excluded.	3	C	100
4. Depending upon the type of neuropsychiatric manifestation, the diagnostic work-up may include lumbar puncture and CSF analysis (primarily to exclude CNS infection), EEG, neuropsychological assessment of cognitive function, consultation with an ophthalmologist, nerve conductional studies and neuroimaging (MRI) to assess nervous system structure and function.*	3	C	100
5. A normal MRI of the CNS does not exclude NP-cSLE.*	3	C	100
6. Cognitive impairment should be tested either in collaboration with a neuropsychologist, or using validated tests for cognitive impairment in cSLE, like the Ped-ANAM.	3	C	100
Treatment recommendations			
1. When neuropsychiatric manifestations are caused by an immune or inflammatory process and non-SLE-related causes are excluded, corticosteroids and immunosuppressive therapy are indicated.	3	C	100
2. Antiepileptic drugs are usually not necessary after a single seizure in the absence of MRI lesions and definite epileptic abnormalities on EEG following recovery from the seizure.*	3	C	100
3. Long-term antiepileptic therapy should be considered for recurrent seizures.*	3	C	93
4. There is a need for paediatric NP-cSLE research regarding treatment.	4	D	100

*This statement is based on the EULAR recommendations for adults with NP-cSLE.[68 123]
Level of evidence: for diagnostic and observational studies: 1A, meta-analysis of cohort studies; 1B, randomised controlled study; 2A, controlled study without randomisation; 2B, quasiexperimental study; and for treatment studies: 1A, meta-analysis of randomised controlled trial; B, based on level 2 or extrapolated from level 1; C, based on level 3 or extrapolated from level 1 or 2; D, based on level 4 or extrapolated from level 3 or 4 expert opinion.[22] Strength of recommendation: A, based on level 1 evidence; 3, descriptive study; 4, expert opinion.[25–27] Agreement indicates percent of experts agreeing on the recommendation during the final voting round of the consensus meeting.
ACR, American College of Rheumatology; CNS, central nervous system; CSF, cerebrospinal fluid; EEG, electroencephalogram; EULAR, European League Against Rheumatism; NP-cSLE, neuropsychiatric childhood-onset systemic lupus erythematosus; NP-SLE, neuropsychiatric systemic lupus erythematosus; Ped-ANAM, Pediatric Automated Neuropsychological Assessment Metrics; SLE, systemic lupus erythematosus.

British Isles Lupus Assessment Group index or the Systemic Lupus Erythematosus Disease Activity Index 2000.[73 75–77] Disease damage should also be comprehensively assessed annually, for example using the paediatric version of the SLICC/American College of Rheumatology Damage Index.[71]

A broad range of ocular manifestations including retinopathy or optic nerve disease can occur in cSLE. Additionally, two of the most commonly used drugs for SLE, corticosteroids and hydroxychloroquine (HCQ), can affect the eyes.[78–80] Therefore, it is recommended that patients have access to the expertise of paediatric ophthalmology. Paucity of evidence regarding ophthalmological risks due to long-term corticosteroids and HCQ use means that annual examination of the eyes should be considered in the paediatric age group.[81]

Despite minimal published evidence supporting the benefits of sun protection in patients with cSLE, sunscreens are widely recommended to prevent photosensitive rashes and as part of general disease management. One study in 11 adult patients with SLE showed that some, but not all types of sunscreen prevented the development of ultraviolet radiation-induced skin lesions.[82]

Adolescent patients need to be supported through the transition process and prepared for transfer of their care to the adult services once they reach adulthood. During adolescence, patients need to develop self-management skills and become responsible for their own health.[83–86] One of the major challenges during the transitional process is non-adherence to treatments,[84 87] which should be addressed frequently at outpatient clinics. EULAR recommendations for this transitional process have been published to support professionals in designing a coordinated transition programme.[88]

General treatment recommendations

It is recommended that all children with lupus should be on HCQ routinely. A systematic review of 95 articles analysing the beneficial and adverse effects of antimalarial therapies such as HCQ in adults with SLE showed a broad spectrum of beneficial effects, such as a higher remission rate, less relapses and less accrual of damage. Additionally, HCQ has a favourable safety profile.[89] Adult studies show that long-term use of HCQ is relatively safe, although the risk of retinopathy increases with the increasing cumulative dose.[89] Unfortunately, no such evidence is available for children with cSLE, but studies in patients with juvenile idiopathic arthritis show that doses up to 6 mg/kg/day (based on lean body weight) are safe to use.[90]

Lack of adherence has been associated with a higher disease activity and more damage.[91–93] Rates of non-adherence can be as high as 50% and disease severity does not guarantee medication adherence.[94] Therefore, adherence should be checked whenever a patient shows poor response to a treatment, measuring medication (trough) levels may be helpful to detect non-adherence. When a patient experiences side effects from a drug, choice of therapy will need to be reassessed and switched if necessary. If disease severity is such that tapering of oral prednisolone is not possible despite adequate compliance to oral prednisone and HCQ, addition of a disease-modifying antirheumatic drug (DMARD) is recommended to improve disease control and permit subsequent corticosteroid tapering. Examples of DMARDs often used include mycophenolate mofetil, azathioprine, methotrexate or cyclophosphamide in severe cases.

The use of rituximab has been described in six studies including a total of 115 individual patients with cSLE. All patients had acute, life-threatening symptoms or symptoms that

did not respond to standard treatment. Two dose regimens were described, which both proved to be effective and safe in the majority of the patients.[95–100]

Diagnostic recommendations

NP-cSLE can be a common manifestation of lupus in children.[101–108] To promote uniformity and comparability between NP-SLE manifestations in children and adults and in view of the limited available evidence in NP-cSLE, it is recommended that the ACR nomenclature and case definitions for NP-SLE[109] are also used in cSLE. It must be taken into account however that the ACR nomenclature was designed for adults and some of the diagnostic or screening tests listed here cannot be used for children. As is the case in adult-onset NP-SLE, no single clinical, laboratory, neuropsychological or imaging test can be used in children to differentiate NP-cSLE from other causes of neuropsychiatric manifestations. There have been some small studies aiming at identification of specific biomarkers or imaging techniques for neuropsychiatric involvement in cSLE, but large controlled studies are lacking.[110–122] Therefore, the recommendation regarding the diagnostic evaluation of neuropsychiatric symptoms is adopted from the adult EULAR recommendations.[68 123]

It is important to make a detailed and thorough assessment of any patient with suspected NP-cSLE. In the context of a suspected NP-cSLE diagnosis or worsening of neuropsychiatric disease, an initial comprehensive work-up should include all other potential underlying causes, including infections, hypertension, metabolic abnormalities or adverse effects of medication. A systematic approach is recommended, with the specific symptoms guiding the type of diagnostic procedure.

Importantly, not all NP-cSLE manifestations can be detected with conventional MRI techniques.[124 125] In addition, conventional MRI techniques (as well as novel MRI imaging modalities) may be unspecific for central nervous system involvement due to cSLE or to other causes. Formal neuropsychiatric testing by a neuropsychologist can be used to ascertain the presence of neurocognitive dysfunction. However, as a neuropsychologist is not always available, a helpful screening tool is the Pediatric Automated Neuropsychological Assessment Metrics, which can be used by non-specialists to screen patients for possible neurocognitive dysfunction[126 127]

Treatment recommendations

The evidence for the treatment of NP-cSLE in children is especially limited. Recommendations are therefore based principally on adult recommendations for the management of NP-SLE,[123] adapted for use in children by the expert panel. It was noted that this remains an important area for future clinical research. When non-SLE-related causes for neuropsychiatric symptoms or signs are excluded, corticosteroids and immunosuppressive therapy are indicated.[123]

Recurrent seizures in SLE may benefit from antiepileptic treatment. However, one single seizure without evidence for epileptic activity on electroencephalogram in the brain is usually not an indication for antiepileptic treatment. Undertaking a careful evaluation seeking and treating the underlying cause, including anti-inflammatory treatment of potential NP-cSLE, most often suffices to prevent further seizures.

DISCUSSION

A total of 35 recommendations for diagnosis, management and treatment for cSLE (25 recommendations) and NP-cSLE (10 recommendations) have been formulated. All recommendations were accepted with >80% agreement.

These recommendations are intended to help specialists with decisions regarding the general care for patients with cSLE. Notably, recommendations regarding the management of nephritis in cSLE and paediatric APS will be published separately.

It must be noted that good quality evidence regarding diagnosis and treatment in cSLE is limited. Due to lack of robust evidence underpinning some statements, the expert panel refrained from being too specific regarding diagnostic procedures, monitoring intervals or specific drug treatments. This emphasises the need for more research on diagnostic procedures, as well as treatment in this population. International collaboration will be vital, as large cohorts are difficult to achieve.

In conclusion, the SHARE project has resulted in recommendations on diagnosis, management and treatment of cSLE and NP-cSLE, based on best available evidence and expert opinion. These recommendations should facilitate the optimisation of the management of this rare disease.

Author affiliations
[1]Wilhelmina Children's Hospital, Utrecht, Netherlands
[2]Sophia Children's Hospital, Erasmus University Medical Centre, Rotterdam, Netherlands
[3]University Children's Hospital Ljubljana, Ljubljana, Slovenia
[4]Necker Hospital, Assistance Publique-Hôpitaux de Paris, Paris, France
[5]Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK
[6]General University Hospital, First Faculty of Medicine, Charles University, Prague, Czech Republic
[7]The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada
[8]Bicêtre Hospital, APHP, University of Paris Sud, Paris, France
[9]Hospital for Children and Adolescents, University of Helsinki, Helsinki, Finland
[10]Alder Hey Children's NHS Foundation Trust, Liverpool, UK
[11]Department of Pediatrics, Hacettepe University, Ankara, Turkey
[12]Università degli Studi di Genova and Istituto Giannina Gaslini, Genoa, Italy
[13]Meir Medical Center, Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel
[14]Institute of Translational Medicine, University of Liverpool, Liverpool, UK

Correction notice This paper has been amended since it was published Online First. Owing to a scripting error, some of the publisher names in the references were replaced with 'BMJ Publishing Group'. This only affected the full text version, not the PDF. We have since corrected these errors and the correct publishers have been inserted into the references.

Contributors MWB and SK are senior authors. NW and BV designed the SHARE initiative. NG and NdG performed the systematic literature review, supervised by MWB and SK. Validity assessment of selected papers was done by MWB, SK, TA, AR, IKP, BBM and CP. Recommendations were formulated by NG, MWB and SK. The expert committee consisted of TA, BBM, PB, PD, IKP, PL, LM, SO, CP, AR, AvR, YU, NW, SK, MWB and SDM; they completed the online surveys and/or participated in the subsequent consensus meetings. NG, NdG, SK and MWB prepared the consensus meetings, and NG and NdG chaired the meetings and took minutes. AR and BF facilitated the consensus procedure using nominal group technique. NG, SK and MWB wrote the manuscript, with contribution and approval of all coauthors.

Competing interests None declared.

Patient consent The study did not involve human subjects.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Bandeira M, Buratti S, Bartoli M, *et al*. Relationship between damage accrual, disease flares and cumulative drug therapies in juvenile-onset systemic lupus erythematosus. *Lupus* 2006;15:515–20.
2 Brunner HI, Gladman DD, Ibañez D, *et al*. Difference in disease features between childhood-onset and adult-onset systemic lupus erythematosus. *Arthritis Rheum* 2008;58:556–62.
3 Hersh AO, von Scheven E, Yazdany J, *et al*. Differences in long-term disease activity and treatment of adult patients with childhood- and adult-onset systemic lupus erythematosus. *Arthritis Rheum* 2009;61:13–20.
4 Livingston B, Bonner A, Pope J. Differences in clinical manifestations between childhood-onset lupus and adult-onset lupus: a meta-analysis. *Lupus* 2011;20:1345–55.

5 Ramírez Gómez LA, Uribe Uribe O, Osio Uribe O, *et al*. Childhood systemic lupus erythematosus in Latin America. the GLADEL experience in 230 children. *Lupus* 2008;17:596–604.
6 Sag E, Tartaglione A, Batu ED, *et al*. Performance of the new SLICC classification criteria in childhood systemic lupus erythematosus: a multicentre study. *Clin Exp Rheumatol* 2014;32:440–4.
7 Fonseca AR, Gaspar-Elsas MI, Land MG, *et al*. Comparison between three systems of classification criteria in juvenile systemic lupus erythematous. *Rheumatology* 2015;54:241–7.
8 Mina R, Brunner HI. Pediatric lupus-are there differences in presentation Response to therapy, and damage accrual compared with adult lupus? rheumatic diseases clinics of North America. *Genetics* 2010;36:53–80. vii-viii.
9 Hiraki LT, Feldman CH, Liu J, *et al*. Prevalence, incidence, and demographics of systemic lupus erythematosus and lupus nephritis from 2000 to 2004 among children in the US Medicaid beneficiary population. *Arthritis Rheum* 2012;64:2669–76.
10 Kamphuis S, Silverman ED. Prevalence and burden of pediatric-onset systemic lupus erythematosus. *Nat Rev Rheumatol* 2010;6:538–46.
11 Pineles D, Valente A, Warren B, *et al*. Worldwide incidence and prevalence of pediatric onset systemic lupus erythematosus. *Lupus* 2011;20:1187–92.
12 Huemer C, Huemer M, Dorner T, *et al*. Incidence of pediatric rheumatic diseases in a regional population in Austria. *J Rheumatol* 2001;28:2116–9.
13 Kaipiainen-Seppänen O, Savolainen A. Incidence of chronic juvenile rheumatic diseases in Finland during 1980-1990. *Clin Exp Rheumatol* 1996;14:441–4.
14 López P, Mozo L, Gutiérrez C, *et al*. Epidemiology of systemic lupus erythematosus in a northern spanish population: gender and age influence on immunological features. *Lupus* 2003;12:860–5.
15 Nightingale AL, Farmer RD, de Vries CS. Incidence of clinically diagnosed systemic lupus erythematosus 1992-1998 using the UK General Practice Research Database. *Pharmacoepidemiol Drug Saf* 2006;15:656–61.
16 Pelkonen PM, Jalanko HJ, Lantto RK, *et al*. Incidence of systemic connective tissue diseases in children: a nationwide prospective study in Finland. *J Rheumatol* 1994;21:2143–6.
17 Series OR. List of rare diseases and synonyms listed in alphabetical order: orphanet, 2016. Available from. http://www.orpha.net/orphacom/cahiers/docs/GB/List_of_rare_diseases_in_alphabetical_order.pdf. (cited 2016 20-01-2017).
18 Wulffraat NM, Vastert B, consortium S; SHARE consortium. Time to share. *Pediatr Rheumatol Online J* 2013;11:5.
19 Enders FB, Bader-Meunier B, Baildam E, *et al*. Consensus-based recommendations for the management of juvenile dermatomyositis. *Ann Rheum Dis* 2017;76:329–40.
20 Giancane G, Ter Haar NM, Wulffraat N, *et al*. Evidence-based recommendations for genetic diagnosis of familial mediterranean fever. *Ann Rheum Dis* 2015;74:635–41.
21 ter Haar NM, Oswald M, Jeyaratnam J, *et al*. Recommendations for the management of autoinflammatory diseases. *Ann Rheum Dis* 2015;74:1636–44.
22 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.
23 Leclercq E, Leeflang MM, van Dalen EC, *et al*. Validation of search filters for identifying pediatric studies in PubMed. *J Pediatr* 2013;162:629–34.
24 T.C Hjptgs. *Cochrane Handbook for Systematic Reviews of Interventions*, 2013.
25 Whiting P, Rutjes AW, Dinnes J, *et al*. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004;8:iii, 1-234.
26 Zhang W, Doherty M, Pascual E, *et al*. EULAR evidence based recommendations for gout. part I: diagnosis. Report of a task force of the standing Committee for International clinical studies including therapeutics (ESCISIT). *Ann Rheum Dis* 2006;65:1301–11.
27 Zhang W, Doherty M, Bardin T, *et al*. EULAR evidence based recommendations for gout. Part II: management. Report of a task force of the EULAR standing Committee for International clinical studies including therapeutics (ESCISIT). *Ann Rheum Dis* 2006;65:1312–24.
28 vdVA DAL. A group process model for problem identification and program planning. *J Appl Behav Sci* 1971;7:466–92.
29 Tarr T, Dérfalvi B, Győri N, *et al*. Similarities and differences between pediatric and adult patients with systemic lupus erythematosus. *Lupus* 2015;24:796–803.
30 Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 19971725;40:1725.
31 Petri M, Orbai AM, Alarcón GS, *et al*. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for Systemic lupus erythematosus. *Arthritis Rheum* 2012;64:2677–86.
32 Abdwani R, Rizvi SG, El-Nour I. Childhood systemic lupus erythematosus in Sultanate of Oman: demographics and clinical analysis. *Lupus* 2008;17:683–6.
33. Bader-Meunier BBM AJ, Haddad E, Cochat P, *et al*. Initial presentation of childhood-onset systemic lupus erythematosus: a French multicenter study. 2005. *J pediatr* 2005:146.
34 Buoncompagni A, Barbano GC, Pistoia V, *et al*. Childhood systemic lupus erythematosus: a review of 30 cases. *Clin Exp Rheumatol* 1991;9:425–30.
35 Chiang LL, Lin YT, Chan HY, *et al*. Differential manifestations of prepubescent, pubescent and postpubescent pediatric patients with systemic lupus erythematosus: a retrospective study of 96 chinese children and adolescents. *Pediatr Rheumatol Online J* 2012;10:12.
36 Hiraki LT, Benseler SM, Tyrrell PN, *et al*. Clinical and laboratory characteristics and long-term outcome of pediatric systemic lupus erythematosus: a longitudinal study. *J Pediatr* 2008;152:550–6.
37 Olowu W. Childhood-onset systemic lupus erythematosus. *J Natl Med Assoc* 2007;99:777–84.
38 Watson L, Leone V, Pilkington C, *et al*. Disease activity, severity, and damage in the UK Juvenile-Onset systemic lupus Erythematosus Cohort. *Arthritis Rheum* 2012;64:2356–65.
39 Campos LM, Kiss MH, Scheinberg MA, *et al*. Antinucleosome antibodies in patients with juvenile systemic lupus erythematosus. *Lupus* 2006;15:496–500.
40 Hinze CH, Suzuki M, Klein-Gitelman M, *et al*. Neutrophil gelatinase-associated lipocalin is a predictor of the course of global and renal childhood-onset systemic lupus erythematosus disease activity. *Arthritis Rheum* 2009;60:2772–81.
41 Jesus AA, Campos LM, Liphaus BL, *et al*. Anti-C1q, anti-chromatin/nucleosome, and anti-dsDNA antibodies in juvenile systemic lupus erythematosus patients. *Rev Bras Reumatol* 2012;52:976–81.
42 Jesus AA, Silva CA, Carneiro-Sampaio M, *et al*. Anti-C1q antibodies in juvenile-onset systemic lupus erythematosus. *Ann N Y Acad Sci* 2009;1173:235–8.
43 Jurencák R, Fritzler M, Tyrrell P, *et al*. Autoantibodies in pediatric systemic lupus erythematosus: ethnic grouping, cluster analysis, and clinical correlations. *J Rheumatol* 2009;36:416–21.
44 Lehman TJ, Hanson V, Singsen BH, *et al*. The role of antibodies directed against double-stranded DNA in the manifestations of systemic lupus erythematosus in childhood. *J Pediatr* 1980;96:657–61.
45 Tang X, Huang Y, Deng W, *et al*. Clinical and serologic correlations and autoantibody clusters in systemic lupus erythematosus: a retrospective review of 917 patients in South China. *Medicine* 2010;89:62–7.
46 Wu FQ, Zhao Q, Cui XD, *et al*. C1q and anti-C1q antibody levels are correlated with disease severity in chinese pediatric systemic lupus erythematosus. *Rheumatol Int* 2011;31:501–5.
47 Wu JF, Yang YH, Wang LC, *et al*. Antinucleosome antibodies correlate with the disease severity in children with systemic lupus erythematosus. *J Autoimmun* 2006;27:119–24.
48 Al-Mayouf SM, Abanomi H, Eldali A. Impact of C1q deficiency on the severity and outcome of childhood systemic lupus erythematosus. *Int J Rheum Dis* 2011;14:81–5.
49 Pickering MC, Botto M, Taylor PR, *et al*. Systemic lupus erythematosus, complement deficiency, and apoptosis. *Adv Immunol* 2000;76:227–324.
50 Crow YJ. Lupus: how much "complexity" is really (just) genetic heterogeneity? *Arthritis Rheum* 2011;63:3661–4.
51 Bader-Meunier B, Cavé H, Jeremiah N, *et al*. Are RASopathies new monogenic predisposing conditions to the development of systemic lupus erythematosus? Case report and systematic review of the literature. *Semin Arthritis Rheum* 2013;43:217–9.
52 Ahmed AM E-SM. Asymptomatic cardiac involvement in children with systemic lupus erythematosus. *J Med Sci* 2006;6:944–9.
53 Al-Abbad AJ, Cabral DA, Sanatani S, Sandor GGS, *et al*. Echocardiography and pulmonary function testing in childhood onset systemic lupus erythematosus. *Lupus* 2001;10:32–7.
54 Cerveri I, Fanfulla F, Ravelli A, *et al*. Pulmonary function in children with systemic lupus erythematosus. *Thorax* 1996;51:424–8.
55 Ciftçi E, Yalçinkaya F, Ince E, *et al*. Pulmonary involvement in childhood-onset systemic lupus erythematosus: a report of five cases. *Rheumatology* 2004;43:587–91.
56 de Jongste JC, Neijens HJ, Duiverman EJ, *et al*. Respiratory tract disease in systemic lupus erythematosus. *Arch Dis Child* 1986;61:478–83.
57 El-Dessoky El Shahawy E MA, Algoubashy AA, Abo-Warda MH, et al. pleuropulmonary manifestations in juvenile onset systemic lupus erythematosus: assessment by pulmonary function tests and multidetector computed tomography. *The Egyptian Rheumatologist* 2011;33:163–9.
58 Gazarian M, Feldman BM, Benson LN, *et al*. Assessment of myocardial perfusion and function in childhood systemic lupus erythematosus. *J Pediatr* 1998;132:109–16.
59 Trapani S, Camiciottoli G, Ermini M, *et al*. Pulmonary involvement in juvenile systemic lupus erythematosus: a study on lung function in patients asymptomatic for respiratory disease. *Lupus* 1998;7:545–50.
60 Beresford MW, Cleary AG, Sills JA, *et al*. Cardio-pulmonary involvement in juvenile systemic lupus erythematosus. *Lupus* 2005;14:152–8.
61 Parodi A, Davì S, Pringe AB, *et al*. Macrophage activation syndrome in juvenile systemic lupus erythematosus: a multinational multicenter study of thirty-eight patients. *Arthritis Rheum* 2009;60:3388–99.
62 Bennett TD, Fluchel M, Hersh AO, *et al*. Macrophage activation syndrome in children with systemic lupus erythematosus and children with juvenile idiopathic arthritis. *Arthritis Rheum* 2012;64:4135–42.

63 Assari R, Ziaee V, Mirmohammadsadeghi A, *et al*. Dynamic changes, Cut-Off Points, sensitivity, and specificity of Laboratory Data to differentiate macrophage activation syndrome from active disease. *Dis Markers* 2015;2015:1–8.
64 Allen CE, Yu X, Kozinetz CA, *et al*. Highly elevated ferritin levels and the diagnosis of hemophagocytic lymphohistiocytosis. *Pediatr Blood Cancer* 2008;50:1227–35.
65 Mina R, Klein-Gitelman MS, Nelson S, *et al*. Validation of the systemic lupus erythematosus responder index for use in juvenile-onset systemic lupus erythematosus. *Ann Rheum Dis* 2014;73:401–6.
66 Mina R, Klein-Gitelman MS, Ravelli A, *et al*. Inactive disease and remission in childhood-onset systemic lupus erythematosus. *Arthritis Care Res* 2012;64:683–93.
67 Brunner HI, Mina R, Pilkington C, *et al*. Preliminary criteria for global flares in childhood-onset systemic lupus erythematosus. *Arthritis Care Res* 2011;63:1213–23.
68 Bertsias G, Ioannidis JP, Boletis J, *et al*. EULAR recommendations for the management of systemic lupus erythematosus. Report of a Task Force of the EULAR standing Committee for International clinical studies including therapeutics. *Ann Rheum Dis* 2008;67:195–205.
69 Bertsias GK, Tektonidou M, Amoura Z, *et al*. Joint European League against Rheumatism and european renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of adult and paediatric lupus nephritis. *Ann Rheum Dis* 2012;71:1771–82.
70 Mosca M, Tani C, Aringer M, *et al*. European League against Rheumatism recommendations for monitoring patients with systemic lupus erythematosus in clinical practice and in observational studies. *Ann Rheum Dis* 2010;69:1269–74.
71 Gutiérrez-Suárez R, Ruperto N, Gastaldi R, *et al*. A proposal for a pediatric version of the systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index based on the analysis of 1,015 patients with juvenile-onset systemic lupus erythematosus. *Arthritis Rheum* 2006;54:2989–96.
72 Rygg M, Pistorio A, Ravelli A, *et al*. A longitudinal PRINTO study on growth and puberty in juvenile systemic lupus erythematosus. *Ann Rheum Dis* 2012;71:511–7.
73 Brunner HI, Higgins GC, Klein-Gitelman MS, *et al*. Minimal clinically important differences of disease activity indices in childhood-onset systemic lupus erythematosus. *Arthritis Care Res* 2010;62:950–9.
74 Brunner HI, Silverman ED, Bombardier C, *et al*. European Consensus Lupus Activity Measurement is sensitive to change in disease activity in childhood-onset systemic lupus erythematosus. *Arthritis Rheum* 2003;49:335–41.
75 Lattanzi B, Consolaro A, Solari N, *et al*. Measures of disease activity and damage in pediatric systemic lupus erythematosus: british Isles Lupus Assessment Group (BILAG), European Consensus Lupus Activity Measurement (ECLAM), Systemic lupus activity measure (SLAM), Systemic lupus Erythematosus Disease Activity Index (SLEDAI), Physician's Global Assessment of Disease Activity (MD Global), and Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SLICC/ACR DI; SDI). *Arthritis Care Res* 2011;63(Suppl 11):112–7.
76 Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol* 2002;29:288–91.
77 Marks SD, Pilkington C, Woo P, *et al*. The use of the British Isles Lupus Assessment Group (BILAG) index as a Valid tool in assessing disease activity in childhood-onset Systemic lupus erythematosus. *Rheumatology* 2004;43:1186–9.
78 Al Mayouf SM AHA. Ocular manifestations of SLE in children. *Saudi Med J* 2003;24:964–6.
79 Hoes JN, Jacobs JW, Boers M, *et al*. EULAR evidence-based recommendations on the management of systemic glucocorticoid therapy in rheumatic diseases. *Ann Rheum Dis* 2007;66:1560–7.
80 Marmor MF, Kellner U, Lai TY, *et al*. Revised recommendations on screening for chloroquine and hydroxychloroquine retinopathy. *Ophthalmology* 2011;118:415–22.
81 ACo R. *Position statement: screening for Hydroxychloroquine Retinopathy*, 2011.
82 Stege H, Budde MA, Grether-Beck S, *et al*. Evaluation of the capacity of sunscreens to photoprotect lupus erythematosus patients by employing the photoprovocation test. *Photodermatol Photoimmunol Photomed* 2000;16:256–9.
83 Tucker LB, Cabral DA. Transition of the adolescent patient with rheumatic disease: issues to consider. *Rheum Dis Clin North Am* 2007;33:661–72.
84 Lawson EF, Hersh AO, Applebaum MA, *et al*. Self-management skills in adolescents with chronic rheumatic disease: a cross-sectional survey. *Pediatr Rheumatol Online J* 2011;9:35.
85 Hersh AO, Pang S, Curran ML, *et al*. The challenges of transferring chronic illness patients to adult care: reflections from pediatric and adult rheumatology at a US academic center. *Pediatr Rheumatol Online J* 2009;7:13.
86 Falcini F, Nacci F. Systemic lupus erythematosus in the young: the importance of a transition clinic. *Lupus* 2007;16:613–7.
87 Felsenstein S, Reiff AO, Ramanathan A. Transition of Care and Health-Related Outcomes in Pediatric-Onset systemic lupus erythematosus. *Arthritis Care Res* 2015;67:1521–8.
88 Foster HE, Minden K, Clemente D, *et al*. EULAR/PReS standards and recommendations for the transitional care of young people with juvenile-onset rheumatic diseases. *Ann Rheum Dis* 2016;76.
89 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.
90 Ziering CL, Rabinowitz LG, Esterly NB. Antimalarials for children: indications, toxicities, and guidelines. *J Am Acad Dermatol* 1993;28(5 Pt 1):764–70.
91 Uribe AG, Alarcón GS, Sanchez ML, *et al*. Systemic lupus erythematosus in three ethnic groups. XVIII. factors predictive of poor compliance with study visits. *Arthritis Rheum* 2004;51:258–63.
92 Rojas-Serrano J, Cardiel MH. Lupus patients in an emergency unit. causes of consultation, hospitalization and outcome. A cohort study. *Lupus* 2000;9:601–6.
93 Koneru S, Kocharla L, Higgins GC, *et al*. Adherence to medications in systemic lupus erythematosus. *J Clin Rheumatol* 2008;14:195–201.
94 M. R. Adherence to Pediatric Medical Regimens. *Handbook of Child Psychology and developmental science*. 7 ed. New York NY: John Wiley & Sons Inc, 2010.
95 Willems M, Haddad E, Niaudet P, *et al*. Rituximab therapy for childhood-onset systemic lupus erythematosus. *J Pediatr* 2006;148:623–7.
96 Polido-Pereira J, Ferreira D, Rodrigues AM, *et al*. Rituximab use in pediatric autoimmune diseases: four case reports. *Ann N Y Acad Sci* 2009;1173:712–20.
97 Podolskaya A, Stadermann M, Pilkington C, *et al*. B cell depletion therapy for 19 patients with refractory systemic lupus erythematosus. *Arch Dis Child* 2008;93:401–6.
98 Nwobi O, Abitbol CL, Chandar J, *et al*. Rituximab therapy for juvenile-onset systemic lupus erythematosus. *Pediatr Nephrol* 2008;23:413–9.
99 Marks SD, Patey S, Brogan PA, *et al*. B lymphocyte depletion therapy in children with refractory systemic lupus erythematosus. *Arthritis Rheum* 2005;52:3168–74.
100 Watson L, Beresford MW, Maynes C, *et al*. The indications, efficacy and adverse events of rituximab in a large cohort of patients with juvenile-onset SLE. *Lupus* 2015;24:10–17.
101 Parikh S, Swaiman KF, Kim Y. Neurologic characteristics of childhood lupus erythematosus. *Pediatr Neurol* 1995;13:198–201.
102 Sibbitt WL, Brandt JR, Johnson CR, *et al*. The incidence and prevalence of neuropsychiatric syndromes in pediatric onset systemic lupus erythematosus. *J Rheumatol* 2002;29:1536–42.
103 Turkel SB, Miller JH, Reiff A. Case series: neuropsychiatric symptoms with pediatric systemic lupus erythematosus. *J Am Acad Child Adolesc Psychiatry* 2001;40:482–5.
104 Yu HH, Lee JH, Wang LC, *et al*. Neuropsychiatric manifestations in pediatric systemic lupus erythematosus: a 20-year study. *Lupus* 2006;15:651–7.
105 Lim LS, Lefebvre A, Benseler S, *et al*. Psychiatric illness of systemic lupus erythematosus in childhood: spectrum of clinically important manifestations. *J Rheumatol* 2013;40:506–12.
106 Loh WF, Hussain IM, Soffiah A, *et al*. Neurological manifestations of children with systemic lupus erythematosus. *Med J Malaysia* 2000;55:459–63.
107 Singh S, Gupta MK, Ahluwalia J, *et al*. Neuropsychiatric manifestations and antiphospholipid antibodies in pediatric onset lupus: 14 years of experience from a tertiary center of North India. *Rheumatol Int* 2009;29:1455–61.
108 Brunner HI, Jones OY, Lovell DJ, *et al*. Lupus headaches in childhood-onset systemic lupus erythematosus: relationship to disease activity as measured by the systemic lupus erythematosus disease activity index (SLEDAI) and disease damage. *Lupus* 2003;12:600–6.
109 The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. *Arthritis Rheum* 1999;42:599–608.
110 Avcin T, Benseler SM, Tyrrell PN, *et al*. A followup study of antiphospholipid antibodies and associated neuropsychiatric manifestations in 137 children with systemic lupus erythematosus. *Arthritis Rheum* 2008;59:206–13.
111 dos Santos MC, Okuda EM, Ronchezel MV, *et al*. Verbal ability impairment in juvenile systemic lupus erythematosus. *Rev Bras Reumatol* 2010;50:362–74.
112 Mortilla M, Ermini M, Nistri M, *et al*. Brain study using magnetic resonance imaging and proton MR spectroscopy in pediatric onset systemic lupus erythematosus. *Clin Exp Rheumatol* 2003;21:129–35.
113 Prismich G, Hilário MO, Len CA, *et al*. Use of single photon emission computed tomography and magnetic resonance to evaluate central nervous system involvement in patients with juvenile systemic lupus erythematosus. *Braz J Med Biol Res* 2002;35:805–10.
114 Reiff A, Miller J, Shaham B, *et al*. Childhood central nervous system lupus; longitudinal assessment using single photon emission computed tomography. *J Rheumatol* 1997;24:2461–5.
115 Russo R, Gilday D, Laxer RM, *et al*. Single photon emission computed tomography scanning in childhood systemic lupus erythematosus. *J Rheumatol* 1998;25:576–82.
116 Szer IS, Miller JH, Rawlings D, *et al*. Cerebral perfusion abnormalities in children with central nervous system manifestations of lupus detected by single photon emission computed tomography. *J Rheumatol* 1993;20:2143–8.
117 Dong J, Li H, Wang JB, *et al*. Predictors for neuropsychiatric development in chinese adolescents with systemic lupus erythematosus. *Rheumatol Int* 2012;32:2681–6.
118 Falcini F, De Cristofaro MT, Ermini M, *et al*. Regional cerebral blood flow in juvenile systemic lupus erythematosus: a prospective SPECT study. single photon emission computed tomography. *J Rheumatol* 1998;25:583–8.
119 Mostafa GA, Ibrahim DH, Shehab AA, *et al*. The role of measurement of serum autoantibodies in prediction of pediatric neuropsychiatric systemic lupus erythematosus. *J Neuroimmunol* 2010;227(1-2):195–201.

120 Mostafa GA, Nazif HK, El-Shahawi HH, *et al*. Antineuronal antibodies and electroneurophysiological studies in pediatric patients with neuropsychiatric systemic lupus erythematosus. *Pediatr Allergy Immunol* 2009;20:192–9.
121 Papero PH, Bluestein HG, White P, *et al*. Neuropsychologic deficits and antineuronal antibodies in pediatric systemic lupus erythematosus. *Clin Exp Rheumatol* 1990;8:417–24.
122 Press J, Palayew K, Laxer RM, *et al*. Antiribosomal P antibodies in pediatric patients with systemic lupus erythematosus and psychosis. *Arthritis Rheum* 1996;39:671–6.
123 Bertsias GK, Ioannidis JP, Aringer M, *et al*. EULAR recommendations for the management of systemic lupus erythematosus with neuropsychiatric manifestations: report of a task force of the EULAR standing committee for clinical affairs. *Ann Rheum Dis* 2010;69:2074–82.
124 Al-Obaidi M, Saunders D, Brown S, *et al*. Evaluation of magnetic resonance imaging abnormalities in juvenile onset neuropsychiatric systemic lupus erythematosus. *Clin Rheumatol* 2016;35:2449–56.
125 Gulati G, Jones JT, Lee G, *et al*. Altered Blood-Brain Barrier Permeability in Patients with systemic lupus erythematosus: a novel Imaging Approach. *Arthritis Care Res* 2017;69:299-305.
126 Brunner HI, Klein-Gitelman MS, Zelko F, *et al*. Validation of the Pediatric Automated Neuropsychological Assessment Metrics in childhood-onset systemic lupus erythematosus. *Arthritis Care Res* 2013;65:372–81.
127 Vega-Fernandez P, Vanderburgh White S, Zelko F, *et al*. Cognitive Performance scores for the Pediatric Automated Neuropsychological Assessment Metrics in Childhood-Onset systemic lupus erythematosus. *Arthritis Care Res* 2015;67:1119–27.

EXTENDED REPORT

Seven-year tolerability profile of glucocorticoids use in early rheumatoid arthritis: data from the ESPOIR cohort

Camille Roubille,[1,2] Nathalie Rincheval,[1,3] Maxime Dougados,[4,5] René-Marc Flipo,[6] Jean-Pierre Daurès,[3] Bernard Combe[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-2016-210135).

[1]Rheumatology Department, Lapeyronie Hospital, Montpellier University, Montpellier, France
[2]Internal Medicine and Hypertension Department, Lapeyronie Hospital, Montpellier University, Montpellier, France
[3]Statistiques, University Institute of Clinical Research, EA2415, Montpellier, France
[4]Paris Descartes University, Rheumatology Department, Cochin Hospital, AP-HP, Paris, France
[5]INSERM (U1153): Clinical Epidemiology and Biostatistics, PRES Sorbonne Paris-Cité, Paris, France
[6]Department of Rheumatology, Roger Salengro Hospital, Lille, France

Correspondence to
Professor Bernard Combe, Département de Rhumatologie, Hôpital Lapeyronie, Montpellier cedex 34295, France; b-combe@chu-montpellier.fr

Received 28 June 2016
Revised 14 January 2017
Accepted 21 January 2017
Published Online First 17 February 2017

► http://dx.doi.org/10.1136/annrheumdis-2017-211187

To cite: Roubille C, Rincheval N, Dougados M, *et al*. *Ann Rheum Dis* 2017;**76**:1797–1802.

ABSTRACT

Objective To explore the 7-year tolerability profile of glucocorticoids (GC) for early rheumatoid arthritis (RA).

Methods We examined data for 602 patients with RA from the early arthritis Etude et Suivi des POlyarthrites Indifférenciées Récentes (ESPOIR) cohort (<6 months disease duration) stratified into two groups: with or without GC treatment at least once during follow-up (median 7 years (IQR 0.038–7.65)). The main outcome was a composite of death, cardiovascular disease (including myocardial ischaemia, cerebrovascular accident and heart failure), severe infection and fracture.

Results Among the 602 patients with RA (476 women (79%), mean age 48±12 years), 386 with GC (64.1%) received low-dose prednisone (mean 3.1±2.9 mg/day for the entire follow-up): 263 started GC during the first 6 months (68%), and the mean duration of total GC treatment was 1057±876 days. As compared with patients without GC (216 (35.9%)), those with GC showed greater use of non-steroidal anti-inflammatory drugs, synthetic and biological disease-modifying antirheumatic drugs and had more active disease disability, higher C reactive protein and anticitrullinated protein antibody levels. Among 65 events (7 deaths, 14 cardiovascular diseases, 19 severe infections and 25 fractures), 44 and 21 occurred in patients with and without GC (p=0.520). Infections were more frequent, although not significantly, in patients with than without GC (p=0.09). On weighted Cox proportional-hazards analysis, with use of propensity score and inverse-probability-of-treatment weighting, and including age, gender, history of hypertension and GC treatment, outcomes did not differ with and without GC (p=0.520; HR=0.889; 95% CI 0.620 to 1.273).

Conclusions This 7-year analysis of the ESPOIR cohort supports the good safety profile of very low-dose GC for early active RA.

INTRODUCTION

Glucocorticoids (GC) are commonly prescribed for patients with rheumatoid arthritis (RA), especially in the early stages of the disease. However, beyond their symptomatic and structural benefits,[1] the risk/benefit ratio of GC remains controversial,[2] with concerns about the tolerability profile. GC adverse events have been reported to be time-dependent and dose-dependent, so international guidelines support the use of the lowest dose for the shortest duration.[3 4]

GC use in RA remains controversial, mostly because of concerns about long-term safety outcomes, including cardiovascular disease (CVD), infection, diabetes, weight gain, osteoporosis and fracture.[5] Nevertheless, despite the fear of adverse events, GC are widely used in RA, especially in early active diseases, with variable dosage and duration. For instance, in the UK, 50% of patients with incident RA were reported to receive GC in primary care, with more than 50% receiving doses >10 mg/day.[6] In the German Course And Prognosis of Early Arthritis (CAPEA) inception cohort, 77% of patients initially received oral GC, 20% receiving low-dose GC (<7.5 mg/day) and 35% high-dose GC (≥20 mg/day).[7] By contrast, in the Canadian CATCH cohort, only 42% of patients started on GC; 48% received oral GC (≤10 mg/day) and 38% intra-articular or intramuscular GC.[8] In another inception cohort from Latin America, 64% of early patients with RA took GC (80% ≤10 mg/day of prednisone).[9] Notably, despite the longstanding use of GC for daily RA management, few strong evidence-based safety data are available.[10 11] Therefore, the safety of GC remains on the research agenda, especially the tolerability profile in early RA.

In the Etude et Suivi des POlyarthrites Indifférenciées Récentes (ESPOIR) cohort (see online supplementary method 1),[12 13] more than half of patients received GC at least once over 5 years after inclusion, especially during the first 6 months of follow-up.[14] Thus, investigation of consecutive events occurring after GC initiation in this cohort may give insight into the long-term safety profile of GC use in a real-life setting.

Here, we aimed to explore the 7-year tolerability profile of GC for patients with recent-onset RA (ESPOIR cohort) by determining the association between GC use and major safety events, including death, CVD, severe infection and fracture.

PATIENTS AND METHODS

Study design and setting

The ESPOIR cohort is a French prospective multicentre observational cohort sponsored by the French Society of Rheumatology that included patients (aged 18–70 years) with early arthritis from 14 rheumatology centres in France. To be included, patients had to have inflammatory arthritis in at least two swollen joints lasting from 6 weeks to 6 months, with the potential to develop

into RA, and be naïve to disease-modifying antirheumatic drugs (DMARDs) and GC therapy. The exclusion criterion was early inflammatory joint disease meeting criteria for a definite diagnosis other than RA or exhibiting features that ruled out progression to RA.[12] The ESPOIR cohort included 813 patients between 2002 and 2005. The objective, design and characteristics of the cohort were previously described[13] (see online supplementary method 1). The database for the present study was locked in 2013 at the 7-year time point. The protocol of the ESPOIR cohort was approved by the ethics committee of Montpellier, France (no. 020307). All patients gave their signed informed consent before inclusion.

Patients and GC use

Among the 813 patients included, we selected the 712 who fulfilled the 2010 American College of Rheumatology/European League Against Rheumatism criteria for RA[15] over the 7 years of follow-up (figure 1). We excluded patients with a history of CVD (including myocardial ischaemia, cerebrovascular accident and heart failure), severe infection or fracture because we anticipated that such patients might have had a different profile for GC prescription (expected to less frequently receive a prescription from rheumatologists) and risk of related side effects (likely greater).

We also excluded patients with missing data for GC treatment and those not followed-up to 1 year (figure 1), to provide sufficient monitoring data and a sufficient duration of follow-up to ensure that the impact of GC was a true effect and not due to chance in a short observation period (see online supplementary method 2).

Hence, we examined data for 602 patients with an RA diagnosis. To consider the group of patients with a history of CVD, severe infections and fractures, we also analysed data for 657 patients comprising the 602 included patients and the 55 patients with such a history, and no missing data (figure 1).

The included patients were then classified into two groups by whether or not they received GC at least once (with or without GC) over the 7 years of follow-up. The group with GC comprised patients who received systemic GC treatment (oral, intramuscular or intravenous) at least once between inclusion and the end of follow-up. The group without GC comprised patients who never took GC between inclusion and the end of follow-up. Patients who received only intra-articular injections of GC or inhaled GC were included in the group without GC,[16] given the minimal systemic spreading.

GC use was defined as use (yes or no) when the first safety event occurred. Doses were calculated as prednisone-equivalent based on accepted standards.[17] For each patient with GC treatment, the mean dosage of GC was calculated by dividing the total quantity of GC by the duration of the entire follow-up.

Outcomes and variables

The primary outcome was a composite of new-onset safety events including all-cause mortality, CVD (myocardial

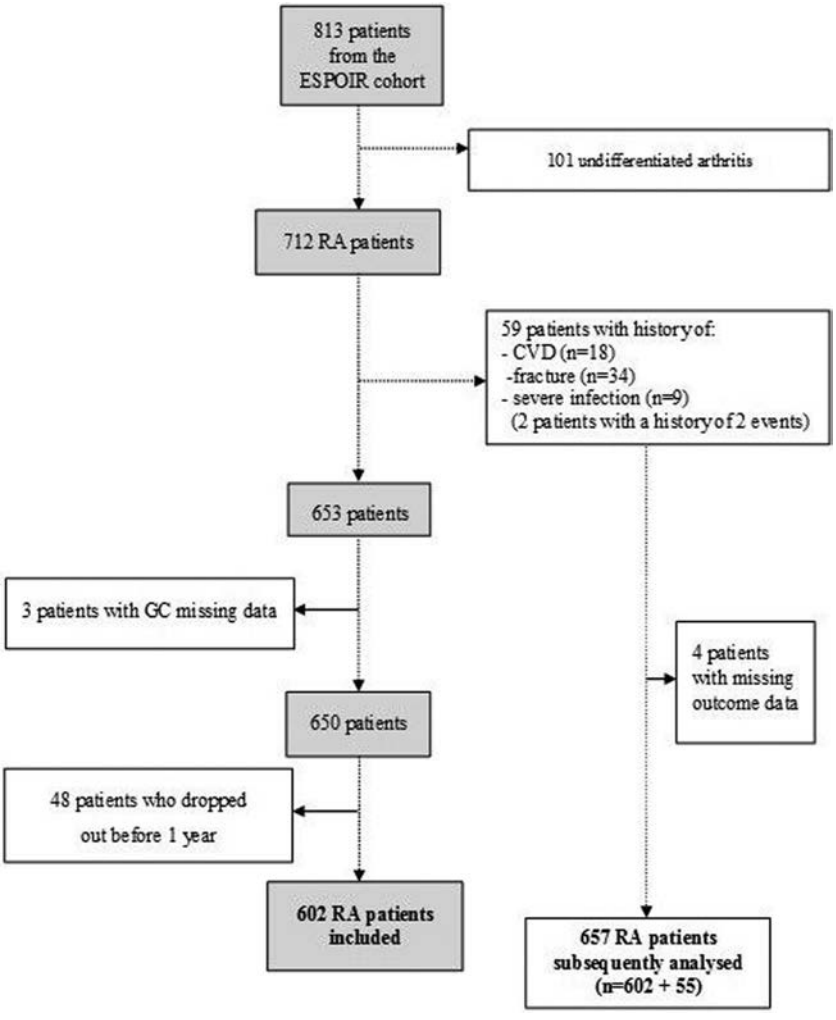


Figure 1 Study design. CVD, cardiovascular disease; ESPOIR, Etude et Suivi des POlyarthrites Indifférenciées Récentes; GC, glucocorticoids; RA, rheumatoid arthritis.

ischaemia, cerebrovascular accident and heart failure), severe infection and fracture. Only new-onset events reported were considered so as to ensure the relevance of the association with GC treatment. We selected as covariates some factors known to affect cardiovascular, infection and fracture risks: age, gender, body mass index (BMI), diabetes, hypertension, smoking status, hypercholesterolaemia and use of non-steroidal anti-inflammatory drugs (NSAIDs). A variable named 'cardiovascular risk' was created to account for the presence of at least one cardiovascular risk factor among hypertension, hypercholesterolaemia, BMI >30 kg/m^2, diabetes and smoking. All all-cause deaths were considered. CVD included myocardial infarction, acute coronary syndrome, angina pectoris, stroke (ischaemic or haemorrhagic) and heart failure. Severe infection was defined as requiring hospitalisation or intravenous antibiotics. Outcomes were recorded in the cohort file.

Statistical analyses

Descriptive statistics are presented as mean±SD or number (%) where appropriate. The non-parametric Mann-Whitney U test was used to compare the distribution of continuous variables and χ^2 test (or Fisher's exact test) to test the association of categorical variables. Continuous variables were transformed into categorical variables with the median or a predetermined threshold. The p values <0.05 were considered significant and all statistical tests were two-sided. The composite primary outcome was compared by χ^2 test (or Fisher's exact test) on univariate analysis. Thereafter, Cox proportional-hazards regression was used to assess the association between GC treatment and outcome, estimating HRs and 95% CIs. To reduce the impact of treatment selection bias and potential confounding, the weighted Cox proportional-hazards model was used with inverse-probability of treatment weighting (IPTW).[18] With this method, weights for patients who had and patients who had not received GC treatment were the inverse of 'propensity score' (PS) and the inverse of '1–PS', respectively. To account for potential confounding by indication, where patients with more severe disease would be more likely to receive GC, we used a PS. The PS is defined as the predicted patient's probability of receiving GC, conditional on a set of observed baseline covariates. The PS was estimated by multiple logistic-regression analysis. Two sets of observed baseline covariates were included in the PS model (see online supplementary figures S1 and S2).[19] The first set of covariates, selected by using the log-rank test, was related to the outcome and the second set, selected using χ^2 test, was related to GC treatment (see online supplementary figures S1 and S2). Of note, all Disease Activity Score in 28 joints (DAS-28) C reactive protein (CRP) levels were not significant, therefore we preferred including each component of the DAS-28 CRP score separately. For both long-rank and χ^2 tests, the level of significance was set at p<0.15. In addition, we used three procedures for selecting variables (forward, backward and stepwise) to obtain the most appropriate logistic-regression equation. All procedures led to the same model. The baseline covariates retained in the final PS regression model were: anticitrullinated protein antibodies (ACPA), diabetes, Health Assessment Questionnaire (HAQ) score, van der Heijde-modified Sharp score (mSHS), cardiovascular risk and patient's overall assessment using visual analogue scale (see online supplementary figure S1). The PS was then included in the Cox proportional-hazards model with the baseline covariates that were significant at 15% on the log-rank test and not already included in the PS model: age, gender, history of hypertension, in addition to GC treatment (yes or no).

In the analysis of the 657 patients, comprising the 602 patients without a history and the 55 patients with a history and sufficient data, the methodology was similar. The two sets of covariates included in the PS model were the same as in the principal analysis, with, in addition, the covariate history of CVD, severe infection or fracture. The baseline covariates retained in the PS regression model for this analysis were the same, except for mSHS (see online supplementary figure S2). The PS was then included in the Cox proportional-hazards model with age, gender, history of hypertension, BMI >30 kg/m^2, CRP, rheumatoid factor (RF), history of CVD, severe infection or fracture, in addition to GC treatment (yes or no). Furthermore, the original studied population could be considered different subgroups on the basis of the cumulative GC dose and total duration of GC treatment over 7 years. These two variables were transformed into a four-level categorical variable by quartiles. First, the log-rank test was used to evaluate the effect of the two covariates with time to event data. Moreover, to adjust for such subgroup differences, an extension of the standard Cox model was used to create two stratified Cox models based on the categorical variable levels. p Value <0.05 was considered to be statistically significant and all statistical tests were two-sided. All statistical analyses involved use of SAS V.9.3 (SAS Institute, Cary, North Carolina, USA).

RESULTS

The study population comprised 602 patients with RA (476 women (79%), mean age 48±12 years; table 1). Mean duration of follow-up was 5.98±1.84 years (median 7 years IQR (0.038–7.65)). Baseline characteristics of the entire sample are shown in table 1: 91% had moderate to high disease activity, and almost 45% were ACPA-positive. A total of 386 (64.1%) received GC during follow-up and 216 never received GC (35.9%). Patients with GC mainly received low-dose prednisone during follow-up (mean 3.1±2.9 mg/day, median 2.4 mg/day (IQR 0.7–5) for the entire follow-up); over half started GC during the first 6 months (n=263, 68.1%). Among the 386 patients who received GC, 73 (19%) received GC for ≤6 months, and 185 (almost 50%) for 2 years (see online supplementary table S1); 69 patients received GC for up to 6 years (18%). The mean duration of total GC treatment was 1057±876 days (median 803 days (IQR 267–1829)) and 280 (72.5%) received GC continuously for longer than 6 months. Six patients received only intravenous or intramuscular GC. Active disease was greater with than without GC, with significantly higher DAS-28-CRP level and HAQ score, as reflected by a greater consumption of DMARDs, biological agents and NSAIDs and higher CRP levels and ACPA titres (table 1).

Effect of GC exposure: composite of death, CVD, severe infection and fracture

A total of 65 events were reported in the entire population: 7 deaths, 14 CVD, 19 severe infections and 25 fractures (table 2).

Deaths occurred between years 4 and 7 and were caused by cancer or malignant blood diseases (n=4), salmonellosis (n=1), ruptured aortic aneurysm (n=1) or unknown cause (n=1).

Among these 65 events, 44 (11.4%) and 21 (9.7%) occurred in patients with and without GC (p=0.520) (table 2). The number of infections was greater, although not significantly, in patients with than without GC (p=0.09).

When we considered patients with events (n=65) and compared patients with and without GC by DAS-28-CRP, those with GC and moderate to high disease activity (DAS-28-CRP score

Table 1 Baseline demographic, clinical, biological and radiographic characteristics of the study population from the ESPOIR cohort (n=602) and those without and with GC

	N	Total study population (n=602)	Without GC (n=216)	With GC (n=386)	p Value*
Age (years)	602	48±12	48.9±11.8	47.5±12.2	0.210
Female	602	476 (79%)	176 (81.5%)	300 (77.7%)	0.277†
BMI, kg/m^2	600	25±4.6	24.9±4.6	25.2±4.6	0.456
BMI >30 kg/m^2	602	83 (13.8%)	28 (13%)	55 (14.3%)	0.643†
Diabetes	602	19 (3.2%)	12 (5.6%)	7 (1.8%)	**0.012**†
Hypertension	602	98 (16.3%)	34 (15.7%)	64 (16.6%)	0.789†
Hypercholesterolaemia	602	84 (14%)	28 (13%)	56 (14.5%)	0.600†
Smokers	602	291 (48.3%)	90 (41.7%)	201 (52.1%)	**0.014**†
Hypertension or hypercholesterolaemia or diabetes or BMI >30 kg/m^2 or smoker	602	401 (66.6%)	133 (61.6%)	268 (69.4%)	**0.050**†
At least one DMARD treatment from baseline to 7 years	567	541 (95%)	178 (93.2%)	363 (96.5%)	**<0.0001**†
At least one biological agent from baseline to 7 years	602	164 (27%)	37 (17.1%)	127 (32.9%)	**<0.0001**†
Consumption of NSAIDs	602	541 (89.9%)	183 (84.7%)	358 (92.7%)	**0.002**†
DAS-28-CRP score	600	4.8±1.2	4.4±1.1	5.0±1.3	**<0.0001**
DAS-28-CRP score	600				**<0.0001**†
≤2.6		22 (3.7%)	13 (6.02%)	9 (2.3%)	
2.6–3.2		32 (5.3%)	13 (6.02%)	19 (4.9%)	
3.2–5.1		324 (54%)	141 (65.3%)	183 (47.7%)	
>5.1		222 (37%)	49 (22.7%)	173 (45.1%)	
HAQ score	602	1.0±0.7	0.8±0.6	1.1±0.7	**<0.0001**
CRP level (mg/L)	602	20.2±33.9	13.2±22.1	24.1±38.5	**<0.0001**
CRP level >10 mg/L	602	274 (45.5%)	76 (35.2%)	198 (51.3%)	**0.0001**†
RF (IU/mL)	602	122.2±445.7	118.7±499.6	124.2±413.1	0.111
IgM-RF positivity	602	317 (52.7%)	107 (49.5%)	210 (54.4%)	0.251†
ACPA titres (IU/mL)	602	555.7±1577.2	416.6±1359	633.5±1683.8	**0.0004**
ACPA positivity	602	268 (44.5%)	76 (35.2%)	192 (49.7%)	**0.0006**†
Typical erosion	571	153 (26.8%)	56 (27.7%)	97 (26.3%)	0.711†
mSHS score	574	5.4±7.8	6.0±8.4	5.1±7.4	0.090

Data are no. (%) or mean±SD.
Bold is related to significant p-Values.
*p Values were assessed by Mann-Whitney U test.
†p Values were assessed by χ^2 test (or Fisher's exact test).
ACPA, anticitrullinated protein antibodies; BMI, body mass index; CRP, C reactive protein; DAS-28-CRP, Disease Activity Score in 28 joints, with C reactive protein level; DMARD, disease-modifying antirheumatic drug; ESPOIR, Etude et Suivi des POlyarthrites Indifférenciées Récentes; GC, glucocorticoids; HAQ, Health Assessment Questionnaire; NSAIDs, non-steroidal anti-inflammatory drugs; RF, rheumatoid factor; mSHS, van der Heijde-modified Sharp score.

>3.2) experienced significantly more events than those without GC (p=0.006). On weighted Cox proportional-hazards analysis (IPTW) (see online supplementary figure S1) including age, gender, history of hypertension and GC treatment, the composite outcome did not differ with and without GC (p=0.520; HR=0.889; 95% CI 0.620 to 1.273). The covariates with significant effect on the composite outcome were age (p=0.02; HR=1.636; 95% CI 1.085 to 2.467) and gender (p=0.003; HR=1.809; 95% CI 1.224 to 2.674).

In the analysis of the 657 patients including those with a history, the Cox IPTW analysis (see online supplementary figure S2) included age, gender, history of hypertension, history of CVD or factures or severe infections, RF, CRP level, BMI >30 kg/m^2 and GC treatment. Again, the composite outcome did not differ with and without GC (p=0.767, HR=0.951; 95% CI 0.684 to 1.323).

Finally, regarding the four-level categorical variable representing cumulative dose and duration of GC treatment, neither the cumulative dose nor duration of GC treatment had an effect on survival (log-rank test, p=0.79 and p=0.57). Also, with two stratified Cox models used to assess the association between the covariates age, gender and history of hypertension and the composite outcome, while controlling for the four-level categorical variable, the results were similar to those with the initial Cox model (see online supplementary figure S3 and table S2).

DISCUSSION

In a cohort of very early RA from a real-life setting monitored for 7 years, we investigated the association between exposure to GC treatment and classical major safety events related to GC (death, CVD, severe infection, fracture). This 7-year data analysis of the ESPOIR cohort did not show any significant difference between patients with RA with and without GC treatment in terms of major safety events. Most of the patients who received GC therapy started GC during the first 6 months and received low-dose therapy. These results support the good safety profile of low-dose GC therapy for early RA and agree with the recent work by Strehl *et al*.[20]

GC are considered a bridging therapy, with short-term symptomatic[21 22] and structural effects. However, the GC risk/benefit balance has little evidence base, with most recent data provided by observational studies. These studies provide the opportunity to explore the real-life tolerability profile of GC, with doses and duration commonly used in daily practice, but often present bias

Table 2 Primary outcome at 7 years (death or cardiovascular disease or severe infection or fracture) in the total sample and patients with and without GC

	Total study population (n=602)	Without GC (n=216)	With GC (n=386)	p Value*
Primary outcome	65 (10.8%)	21 (9.7%)	44 (11.4%)	0.520
Death	7	1	6	0.430
Cardiovascular disease	14	3	11	0.400
Coronary artery disease	8	2	6	–
Stroke	5	1	4	–
Heart failure	1	0	1	–
Severe infections	19	3	16	0.090
Pneumonia	4	2	2	–
Urinary tract infection	7	0	7	–
Digestive	3	1	2	–
Cutaneous	3	0	3	–
Other	2	0	2	–
Fractures	25	14	11	0.150

Data are number or no. (%).
*p Values were assessed by χ^2 test (or Fisher's exact test).
GC, glucocorticoids.

such as confounding by indication.[20] As well, the evidence from randomised clinical trials is scarce.[11 23]

The GC tolerability profile has been reported to depend both on the duration of exposure and dose.[5] Indeed, in addition to a better tolerability profile of a low-dose than high-dose regimen,[24 25] long-term use of low-dose GC has been associated with increased mortality as compared with shorter exposures.[26] Most notably, two recent studies suggested a dose-dependent increase in mortality in RA[27 28]; del Rincón *et al*[27] revealed a daily threshold dose of 8 mg at which all-cause mortality increased with GC dose (adjusted HR=1.78; 95% CI 1.22 to 2.60), and in the German register Rheumatoid Arthritis oBservation of BIologic Therapy (RABBIT), use of GC >5 mg/day was associated with increased mortality risk, independent of RA activity.[28] Moreover, a 10-year follow-up study examined cardiovascular events and deaths in early patients with RA with no history of CVD who were included in a recent open-label randomised trial of low-dose prednisolone (7.5 mg/day) over the first 2 years of early RA Better Anti-Rheumatic PharmacOTherapy (BARFOT+): low-dose prednisolone use was associated with increased incidence of cerebrovascular events and, although not significant, increased mortality.[29] Long-term follow-up of Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA II) patients with early RA who received prednisone at 10 mg/day for at least 2 years revealed increased cardiovascular risk and, although not significant, increased mortality.[30]

In our study, most of the patients who took GC received low-dose GC, <5 mg/day (mean dosage during the entire follow-up 3.1±2.9 mg/day), for which the literature supports an acceptable safety profile.[31]

Cardiovascular tolerance of GC remains controversial. In one meta-analysis of observational studies, GC usage was associated with increased risk of all cardiovascular events (relative risk=1.47; 95% CI 1.34 to 1.60), including myocardial infarction, heart failure and stroke.[32] In another systematic literature review, low-dose GC (<10 mg/day) was associated with major cardiovascular events in four of six studies.[33]

GC therapy has been associated with increased risk of severe infections.[34–36] One systematic review noted the paucity of data on the association between low-dose GC (<10 mg/day prednisone) and risk of infection.[37] In one recent study evaluating patients with RA aged >65 years, the risk of serious infection was increased 30%, 46% or 100% with 5 mg prednisolone used continuously for the last 3 or 6 months or 3 years, respectively, as compared with no use.[38] The increased risk of severe infection was also proportional to the cumulative dosage over 2–3 years.

Potential limitations of the present study are those inherent to observational studies, with potential confounders that could not be taken into account. Moreover, as in many cohort studies, the data are mostly declarative. The events and comorbid diseases were reported by patients, and a potential recall bias cannot be excluded. HAQ and disease activity variables that were selected by the logistic regression analysis and included in the PS were baseline variables, which could also be considered a limitation, because the relationship between GC-related events and disease activity evolution over time could not be evaluated. Using a composite end point is controversial, mostly because it may emphasise each patient's first outcome. Nevertheless, GC toxicity is multifaceted and this way of assessing the most important adverse events may help identify the net effect of GC. We also decided to use a composite outcome including the four most relevant adverse events related to GC treatment, mainly to increase the number of events and to cover the four most worrisome adverse effects of GC therapy defined by a panel of rheumatologists.[39] Moreover, we could not perform a dose-response analysis because of the low doses the patients received. Finally, the relatively small number of incident events might have implied relatively low power.

The present study has many strengths. First, the ESPOIR cohort offered a unique opportunity to explore the long-term impact of GC in very early RA, in a real-life scenario. Importantly, all participants who received GC started treatment after entering the cohort, and only incident safety events were considered. To our knowledge, this is the first cohort study specifically designed to assess (among other data) GC adverse effects and to report the long-term tolerability profile of GC use in early RA. Second, the present 7-year duration of follow-up provides a sufficient period of observation to ensure a true association between GC therapy and adverse effects related to long-term treatment. Third, the ESPOIR cohort has inherent qualities, including the prospective independent collection of data and low rate of missing data and dropout.[12]

Finally, using a PS in this study limited the indication bias, thus reinforcing the association between GC and long-term adverse events. Indeed, several factors such as RA disease activity and comorbidities or history may have an impact on GC prescription. For instance, in one recent observational study, the association between GC use and increased incidence of CVD was negated after adjustment for disease activity and severity, which suggests an effect confounded by indication due to high disease activity.[40] The propensity for prescribing GC is important to consider when evaluating the association between GC use and related outcomes, given that patients with RA with more active disease might have more likely received GC, and conversely, those with history of CVD or severe infections or fractures might have less frequently received GC. Therefore, using a PS allowed for adjusting of patterns that are difficult to fully account for by adjusting for only general and RA-related characteristics in regression modelling. In the ESPOIR cohort, several characteristics differed between patients with and without GC, which highlights the value of using the PS in evaluating the relationship between GC use and outcomes.

CONCLUSIONS

This 7-year data analysis of the ESPOIR cohort did not show any significant difference in major safety events among patients with RA with and without GC treatment. These data support the good safety profile of very low-dose GC therapy in early RA. Although our findings need further confirmation, they strongly support the current recommendations[3] that GC should be used for early RA, with DMARDs, for the shortest period and at the lowest possible dose.

Acknowledgements The authors thank Mohamed Belkacemi for statistical advice.

Contributors All authors participated in the conception and design of the study, contributed to the acquisition of data, participated in the analysis and interpretation of data and read, revised and approved the final manuscript. NR and JPD performed and are responsible for the statistical analyses.

Funding CR received a grant (2838) from the French Society of Rheumatology (SFR). The Etude et Suivi des POlyarthrites Indifférenciées Récentes (ESPOIR) cohort was supported by an unrestricted grant from Merck Sharp and Dohme for the first 5 years. Part of the biological database was supported by two additional grants from Institut National de la Santé et de la Recherche Médicale (INSERM), and the ESPOIR cohort study was also supported by the French Society of Rheumatology, AbbVie, Pfizer and Roche Chugai.

Competing interests None declared.

Ethics approval EC Montpellier University Hospital, France.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Kirwan JR, Bijlsma JW, Boers M, *et al*. Effects of glucocorticoids on radiological progression in rheumatoid arthritis. *Cochrane Database Syst Rev* 2007;(1):CD006356.
2 Kavanaugh A, Wells AF. Benefits and risks of low-dose glucocorticoid treatment in the patient with rheumatoid arthritis. *Rheumatology (Oxford)* 2014;53:1742–51.
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: 2013 update. *Ann Rheum Dis* 2014;73:492–509.
4 Singh JA, Saag KG, Bridges SL Jr, *et al*. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Ann Rheum Dis* 2016;68:1–26.
5 Ethgen O, de Lemos Esteves F, Bruyere O, *et al*. What do we know about the safety of corticosteroids in rheumatoid arthritis? *Curr Med Res Opin* 2013;29:1147–60.
6 Black RJ, Joseph RM, Brown B, *et al*. Half of U.K. patients with rheumatoid arthritis are prescribed oral glucocorticoid therapy in primary care: a retrospective drug utilisation study. *Arthritis Res Ther* 2015;17:375.
7 Albrecht K, Callhoff J, Schneider M, *et al*. High variability in glucocorticoid starting doses in patients with rheumatoid arthritis: observational data from an early arthritis cohort. *Rheumatol Int* 2015;35:1377–84.
8 McKeown E, Bykerk VP, De Leon F, *et al*. Quality assurance study of the use of preventative therapies in glucocorticoid-induced osteoporosis in early inflammatory arthritis: results from the CATCH cohort. *Rheumatology (Oxford)* 2012;51:1662–9.
9 Cardiel MH, Pons-Estel BA, Sacnun MP, *et al*. Treatment of early rheumatoid arthritis in a multinational inception cohort of Latin American patients: the GLADAR experience. *J Clin Rheumatol* 2012;18:327–35.
10 Gaujoux-Viala C, Gossec L. When and for how long should glucocorticoids be used in rheumatoid arthritis? International guidelines and recommendations. *Ann N Y Acad Sci* 2014;1318:32–40.
11 Santiago T, da Silva JA. Safety of glucocorticoids in rheumatoid arthritis: evidence from recent clinical trials. *Neuroimmunomodulation* 2015;22:57–65.
12 Combe B, Rincheval N. Early lessons from the recent-onset rheumatoid arthritis cohort ESPOIR. *Joint Bone Spine* 2015;82:13–17.
13 Combe B, Benessiano J, Berenbaum F, *et al*. The ESPOIR cohort: a ten-year follow-up of early arthritis in France: methodology and baseline characteristics of the 813 included patients. *Joint Bone Spine* 2007;74:440–5.
14 Combe B, Rincheval N, Benessiano J, *et al*. Five-year favorable outcome of patients with early rheumatoid arthritis in the 2000s: data from the ESPOIR cohort. *J Rheumatol* 2013;40:1650–7.
15 Aletaha D, Neogi T, Silman AJ, III, *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.
16 Aviña-Zubieta JA, Abrahamowicz M, De Vera MA, *et al*. Immediate and past cumulative effects of oral glucocorticoids on the risk of acute myocardial infarction in rheumatoid arthritis: a population-based study. *Rheumatology (Oxford)* 2013;52:68–75.
17 Buttgereit F, da Silva JA, Boers M, *et al*. Standardised nomenclature for glucocorticoid dosages and glucocorticoid treatment regimens: current questions and tentative answers in rheumatology. *Ann Rheum Dis* 2002;61:718–22.
18 Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550–60.
19 Brookhart MA, Schneeweiss S, Rothman KJ, *et al*. Variable selection for propensity score models. *Am J Epidemiol* 2006;163:1149–56.
20 Strehl C, Bijlsma JW, de Wit M, *et al*. Defining conditions where long-term glucocorticoid treatment has an acceptably low level of harm to facilitate implementation of existing recommendations: viewpoints from an EULAR task force. *Ann Rheum Dis* 2016;75:952–7.
21 Alten R, Grahn A, Holt RJ, *et al*. Delayed-release prednisone improves fatigue and health-related quality of life: findings from the CAPRA-2 double-blind randomised study in rheumatoid arthritis. *RMD Open* 2015;1:e000134.
22 Cutolo M. Glucocorticoids and chronotherapy in rheumatoid arthritis. *RMD Open* 2016;2:e000203.
23 Bakker MF, Jacobs JW, Welsing PM, *et al*. Low-dose prednisone inclusion in a methotrexate-based, tight control strategy for early rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2012;156:329–39.
24 Da Silva JA, Jacobs JW, Kirwan JR, *et al*. Safety of low dose glucocorticoid treatment in rheumatoid arthritis: published evidence and prospective trial data. *Ann Rheum Dis* 2006;65:285–93.
25 Hoes JN, Jacobs JW, Verstappen SM, *et al*. Adverse events of low- to medium-dose oral glucocorticoids in inflammatory diseases: a meta-analysis. *Ann Rheum Dis* 2009;68:1833–8.
26 Sihvonen S, Korpela M, Mustonen J, *et al*. Mortality in patients with rheumatoid arthritis treated with low-dose oral glucocorticoids. A population-based cohort study. *J Rheumatol* 2006;33:1740–6.
27 del Rincón I, Battafarano DF, Restrepo JF, *et al*. Glucocorticoid dose thresholds associated with all-cause and cardiovascular mortality in rheumatoid arthritis. *Arthritis Rheum* 2014;66:264–72.
28 Listing J, Kekow J, Manger B, *et al*. Mortality in rheumatoid arthritis: the impact of disease activity, treatment with glucocorticoids, TNFα inhibitors and rituximab. *Ann Rheum Dis* 2015;74:415–21.
29 Ajeganova S, Svensson B, Hafström I, BARFOT Study Group. Low-dose prednisolone treatment of early rheumatoid arthritis and late cardiovascular outcome and survival: 10-year follow-up of a 2-year randomised trial. *BMJ Open* 2014;4:e004259.
30 de Hair M, IJff N, Jacobs J, *et al*. Long-term adverse events after daily concomitant treatment with 10 mg prednisone in the 2-year computer assisted management in early rheumatoid arthritis trial-II [abstract]. *Arthritis Rheumatol* 2015;67(Suppl 10):856–7.
31 Caporali R, Todoerti M, Scirè CA, *et al*. Oral low-dose glucocorticoids should be included in any recommendation for the use of non-biologic and biologic disease-modifying antirheumatic drugs in the treatment of rheumatoid arthritis. *Neuroimmunomodulation* 2015;22:104–11.
32 Roubille C, Richer V, Starnino T, *et al*. The effects of tumour necrosis factor inhibitors, methotrexate, non-steroidal anti-inflammatory drugs and corticosteroids on cardiovascular events in rheumatoid arthritis, psoriasis and psoriatic arthritis: a systematic review and meta-analysis. *Ann Rheum Dis* 2015;74:480–9.
33 Ruyssen-Witrand A, Fautrel B, Saraux A, *et al*. Cardiovascular risk induced by low-dose corticosteroids in rheumatoid arthritis: a systematic literature review. *Joint Bone Spine* 2011;78:23–30.
34 Wolfe F, Caplan L, Michaud K. Treatment for rheumatoid arthritis and the risk of hospitalization for pneumonia: associations with prednisone, disease-modifying antirheumatic drugs, and anti-tumor necrosis factor therapy. *Arthritis Rheum* 2006;54:628–34.
35 Lacaille D, Guh DP, Abrahamowicz M, *et al*. Use of nonbiologic disease-modifying antirheumatic drugs and risk of infection in patients with rheumatoid arthritis. *Arthritis Rheum* 2008;59:1074–81.
36 Haraoui B, Jovaisas A, Bensen WG, *et al*. Use of corticosteroids in patients with rheumatoid arthritis treated with infliximab: treatment implications based on a real-world Canadian population. *RMD Open* 2015;1:e000078.
37 Ruyssen-Witrand A, Fautrel B, Saraux A, *et al*. Infections induced by low-dose corticosteroids in rheumatoid arthritis: a systematic literature review. *Joint Bone Spine* 2010;77:246–51.
38 Dixon WG, Abrahamowicz M, Beauchamp ME, *et al*. Immediate and delayed impact of oral glucocorticoid therapy on risk of serious infection in older patients with rheumatoid arthritis: a nested case-control analysis. *Ann Rheum Dis* 2012;71:1128–33.
39 van der Goes MC, Jacobs JW, Boers M, *et al*. Patient and rheumatologist perspectives on glucocorticoids: an exercise to improve the implementation of the European League Against Rheumatism (EULAR) recommendations on the management of systemic glucocorticoid therapy in rheumatic diseases. *Ann Rheum Dis* 2010;69:1015–21.
40 van Sijl AM, Boers M, Voskuyl AE, *et al*. Confounding by indication probably distorts the relationship between steroid use and cardiovascular disease in rheumatoid arthritis: results from a prospective cohort study. *PLoS ONE* 2014;9:e87965.

EXTENDED REPORT

Infections and respiratory tract disease as risk factors for idiopathic inflammatory myopathies: a population-based case–control study

John Svensson,[1,2] Marie Holmqvist,[2,3] Ingrid E Lundberg,[1,3] Elizabeth V Arkema[2]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211174).

[1]Rheumatology unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
[2]Clinical Epidemiology Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
[3]Rheumatology unit, Karolinska University Hospital, Stockholm, Sweden

Correspondence to
John Svensson, Karolinska University Hospital, 171 76, Stockholm, Sweden; john.svensson@ki.se

Received 20 January 2017
Accepted
Published Online First
30 August 2017

ABSTRACT

Objectives To investigate the association between infection or respiratory tract disease and future risk of developing idiopathic inflammatory myopathy (IIM).

Methods A case–control study was performed using Swedish nationwide registers. Adults with newly diagnosed IIM were identified (2002–2011) from the National Patient Register (NPR) and the Swedish Rheumatology Register (n=957). Controls were matched by age, sex and place of residence (n=9476). Outpatient visits and hospitalisations preceding IIM diagnosis indicating infection or respiratory disease were identified from NPR. Conditional logistic regression models were used to calculate OR and 95% CI. Sensitivity analyses were performed by varying the exposure definition, adjusting for previous healthcare consumption and excluding individuals with connective tissue disease, IIM lung phenotype or IIM-associated cancer.

Results Preceding infections were more common in IIM cases compared with controls (13% vs 9%) and were associated with an increased risk of IIM (OR 1.5, 95% CI 1.2 to 1.9). Gastrointestinal and respiratory tract infections were associated with an increased risk of IIM while cutaneous infections were not. Preceding respiratory tract disease was present in 10% of IIM cases and 4% of controls (OR 2.3, 95% CI 1.8 to 3.0). Both upper and lower respiratory tract diseases were associated with an increased risk of IIM. Variations in exposure and outcome definitions did not greatly affect the results.

Conclusions Infections and respiratory tract diseases are associated with an increased risk of IIM which suggests that the triggering of the immune system may take place outside the skeletal muscle.

INTRODUCTION

Idiopathic inflammatory myopathies (IIM) are a heterogeneous group of rare autoimmune disorders characterised by weakness and inflammation in skeletal muscles. Based on clinical and muscle biopsy findings, IIM is often divided into three clinical subtypes, polymyositis (PM), dermatomyositis (DM) and inclusion body myositis (IBM).[1]

IIM is believed to be caused by an interaction between environmental and genetic risk factors. The HLA-DRB1*0301 gene has been identified as the strongest genetic risk factor.[2] Few environmental risk factors have been identified, although ultraviolet light and vitamin D have been discussed.[3 4] Infections could trigger autoimmunity through multiple mechanisms including molecular mimicry[5] and epitope spreading[6] and have been implicated in the development of other rheumatic diseases such as rheumatoid arthritis (RA).[7 8] For IIM, there are case reports and case series showing an increased frequency of infections preceding IIM but population-based studies investigating infection as a risk factor for IIM are lacking.[9 10]

Because the lung is a common extra-muscular manifestation in IIM and many patients display lung involvement at diagnosis,[11] it is of special interest in IIM aetiology. Infectious initiation of inflammation is one suggested mechanism of autoimmunity but the location of the inflammation could be key. A possible site of autoimmunity initiation is the lung's mucosal tissue, which is the first line of defence against some exposures such as infection and smoking.[12] Evidence has accumulated that smoking is a triggering factor for the development of anti-citrullinated peptide antibodies in RA.[13] In IIM, anti-Jo-1 antibody positive cases are more likely to be smokers, which suggests that smoking and the lung play important roles in IIM development as well.[14] A recent case–control study showed that a history of lung disease (sarcoidosis, pneumonia or tuberculosis) reported via questionnaire was associated with IIM, providing further evidence that IIM could start in the lung.[15]

Both infections and respiratory disease could cause disease through different or overlapping mechanisms but it is unclear whether it is the cause of inflammation or the site where it occurs that is important in immune-system activation. We, therefore, aimed to investigate the association between an infection or respiratory tract disease and future risk of developing IIM.

METHODS

Study design

We conducted a population-based case–control study including newly diagnosed cases with IIM and matched controls from the general population. Respiratory tract disease and infections were identified from the time period before IIM diagnosis.

Setting

In Sweden, adult patients with IIM are treated by specialists in internal medicine, rheumatology, dermatology and neurology. There is universal access to publicly funded healthcare, including inpatient, non-primary outpatient and primary care, for all residents. Using the unique personal identification number issued to all Swedish residents, data on

To cite: Svensson J, Holmqvist M, Lundberg IE, *et al*. *Ann Rheum Dis* 2017;**76**:1803–1808.

demographics, morbidity and mortality from national administrative and clinical registers can be linked.

Study population

The National Patient Register (NPR) contains data on hospitalisations since 1987 and outpatient visits since 2001, listing main and up to 10 contributory diagnoses at each visit. The coverage is 99% for hospitalisations and 80% for outpatient care mainly due to lower reporting from private and psychiatric care. As 95% of Swedish specialists taking care of patients with myositis are hospital based, the coverage of outpatient visits is high. Cases with a first ever visit (index date) listing an International Classification of Diseases (ICD) code for IIM were identified from the NPR (see online supplementary table 1). They were required to have ≥2 visits listing an IIM code in a specialist clinic (rheumatology, internal medicine, dermatology or neurology). To exclude possible miss-coding, a subsequent visit to a specialist clinic between 1–12 months from index date was required to be included in the study.[16] Cases were also identified from the Swedish Rheumatology Quality register which includes information provided by rheumatologists on diagnosis, treatment and disease activity including IIM since 2003.

Both register sources were available until December 31 2012 and cases were identified between January 1 2002 and December 31 2011. This allows for a 12-month washout period before study start to exclude prevalent cases and 12 months following index date to allow time for a second visit for individuals identified in the NPR. Because diagnosis is set by name by a rheumatologist in SRQ rather than by ICD code, IIM subdiagnosis was assigned with the following priority: (1) SRQ-given diagnosis, (2) latest diagnosis at a rheumatology clinic. As there is no ICD code available for IBM, the code for PM is commonly used in clinical practice. Therefore, DM is presented separately and PM and IBM are presented together as 'other IIM'.

Controls were identified from the National Population Register and were matched 10:1 to each individual with IIM on age, sex and place of residence. Controls were required to live in Sweden and to have no history of IIM at the time the matched case was diagnosed (index date).

Identification and classification of exposure

The NPR was used to identify the most recent hospitalisation or specialist outpatient visit from 2001 to index date using ICD codes (see online supplementary table 1) indicating infections or respiratory tract disease (non-infectious). In addition, tuberculosis was identified from the Swedish Tuberculosis Register, to which reporting is mandatory from treating physicians and tuberculosis laboratories. To minimise the risk of reversed causality, a latency period of 1 year between exposure and diagnosis was used (hospitalisations or visits indicating infection of respiratory tract disease in the year before IIM diagnosis were not considered as exposures).

Exposures were classified based on anatomical location: gastrointestinal, skin, or respiratory tract for infections and lower or upper airways for respiratory tract diseases (see online supplementary table 1). Groups were analysed separately and were not mutually exclusive.

Other variables

Sex and age were available from the Total Population Register. Previous healthcare consumption was assessed by counting the number of hospitalisations (excluding exposure-related, pregnancy-related and IIM-related visits) occurring within 5 years prior to exposure or within 1-6 years prior to index date for unexposed individuals to account for the 1-year latency period. We chose to identify visits in this time period to allow all individuals to have the same possibility to present and to exclude visits in close proximity to diagnosis as they could be caused by the outcome.

History of a connective tissue disorder (CTD) was defined as having ≥1 hospitalisation or outpatient indicating CTD before the index date. IIM-associated cancer was defined as having a registration of a malignant cancer within 3 years before or after index date. Cancers were identified from the National Cancer Register to which reporting of cancers is mandatory by clinicians and pathologists and coverage is close to 99% of all malignant cancers in Sweden.[17] Having a hospitalisation or outpatient visit indicating fibrosis or lung infiltrates within 3 years before or after the index date was defined as having an IIM lung phenotype.

Statistical analyses

Conditional logistic regression models were used to estimate OR and 95% CI which approximate the relative risk between the exposures and IIM. To investigate if there was a dose–response relationship between exposure and outcome, the number of visits indicating a specific exposure was categorised as 0, 1, 2–4 and ≥5 and added to the model as a single categorical exposure variable.

We performed several sensitivity analyses to test our definition of exposure. First, we investigated whether varying the latency period from 1 year to 0 or 3 years changed the results. Second, we examined a stricter definition of exposure requiring the diagnosis of infection or respiratory tract disease to be listed as the main diagnosis in the NPR. Last, we explored whether the effect differed by varying register sources and time periods used to identify exposures. Only hospitalisations were used to identify more severe infections and respiratory tract disease in one analysis. Because hospitalisation data were available from 1987, we could identify exposures occurring up to 24 years before disease diagnosis. We also used the Prescription Drug Register (PDR), listing all prescribed drug dispensations in Sweden from July 2005, in addition to the NPR. The addition of the PDR enabled us to identify less severe types of exposure treated in primary care where prescription drugs were used (see online supplementary table 1).

Because future IIM patients might have an impaired immune system which can potentially cause poorer health leading to more contact with care providers, previous healthcare consumption was assessed as a proxy for general health, categorised into quartiles and added to the model as a confounding variable in a sensitivity analysis.

To investigate if the associations were primarily caused by other IIM-related conditions, individuals with a history of CTD, IIM-associated cancer or IIM lung phenotype were excluded in a sensitivity analysis.

The study was approved by the Ethics Committee at Karolinska Institutet.

RESULTS

We identified 957 IIM cases and 9476 controls between 2002 and 2011. For cases, mean time since most recent infection and respiratory tract disease was 3.4 and 3.0 years, respectively. The time since infection and respiratory tract disease was similar for controls (table 1).

The most common infections were pneumonia, gastrointestinal and fungal infections while chronic obstructive pulmonary

Table 1 Descriptive characteristics of 957 identified incident IIM cases and matched controls

	IIM cases	Controls
N	957	9476
Women, n (%)	546 (57)	5401 (57)
Age, mean (SD)	59 (15)	59 (15)
Education, n (%)		
>12 years	255 (27)	2554 (27)
10–12 years	392 (41)	3886 (41)
<9 years	297 (31)	2902 (31)
Missing	13 (1)	134 (1)
Diagnosis, n (%)		
Dermatomyositis	301 (31)	
Other IIM*	656 (69)	
History of hospitalisation or outpatient visit indicating exposure†, n (%)		
Infections	125 (13)	877 (9)
Respiratory tract disease	92 (10)	423 (4)
Years since last exposure†, mean (SD)		
Infections	3.4 (2.1)	3.6 (2.3)
Respiratory tract disease	3.0 (2.0)	3.4 (2.2)

*Other IIM includes polymyositis, inclusion body myositis and unspecified myositis
†12-month latency period between exposure and outcome.

disease (COPD), asthma and ILD were the most common respiratory tract diseases (see **online** supplementary table 2).

Previous infections were present in 125 (13%) of IIM cases and 877 (9%) of controls (OR 1.5, 95% CI 1.2 to 1.9). Both gastrointestinal infections (OR 1.9, 95% CI 1.1 to 3.5) and respiratory tract infections (OR 1.6, 95% CI 1.1 to 2.3) were associated with an increased future risk of IIM while skin infections were not (OR 1.2, 95% CI 0.8 to 2.0) (figure 1).

A history of respiratory tract disease was more common in IIM cases compared with controls (n=92, 10% vs n=423, 4%) and was associated with an increased risk of IIM (OR 2.3, 95% CI 1.8 to 3.0). Both lower (OR 2.4, 95% CI 1.8 to 3.0) and upper (OR 1.9, 95% CI 1.3 to 2.7) respiratory tract disease were positively associated with IIM (figure 1).

Individuals with more registered visits indicating infections and respiratory tract disease were at an increased risk to develop IIM (figure 2).

No differences were seen when stratifying by diagnosis for infection (DM OR=1.6 and other IIM OR=1.5) or for respiratory tract disease (DM OR=2.3 and other IIM OR=2.3).

Sensitivity analyses

When the latency period between exposure and outcome was removed, the ORs increased for both infections (1.9 vs 1.5) and respiratory tract disease (2.8 vs 2.1). However, when a latency period of 3 years was used, the ORs remained similar to the main analyses. ORs were similar when requiring the diagnosis of infection or respiratory tract disease to be listed as the main diagnosis in the NPR (table 2).

Using different data sources to identify exposure resulted in similar OR estimates (online supplementary table 3).

When adjusting for previous healthcare consumption, the associations decreased but remained significant for both infections (OR 1.3, 95% CI 1.1 to 1.7) as well as for respiratory tract disease (OR 2.2, 95% CI 1.7 to 2.8).

One-third (n=297, 31%) of the cases were removed when individuals with CTDs (n=144, 15%), IIM-associated cancer (n=106, 11%) or IIM lung phenotype (n=84, 9%) were excluded. A positive association remained for both infections (OR 1.4, 95% CI 1.1 to 1.9) and respiratory tract disease (OR 2.2, 95% CI 1.6 to 3.1).

DISCUSSION

In this nationwide population-based case–control study including almost 1000 IIM cases, individuals with a history of an infection had a 50% increased risk of IIM while respiratory tract disease more than doubled the risk. Investigating exposures occurring more than 1 year before the IIM diagnosis enabled us to examine the potential role of exposure long before disease onset which could have a greater influence on disease development rather than exposures occurring in close proximity to disease onset.

Most studies suggesting infections as risk factors for IIM are based on case reports and have examined events in close proximity to diagnosis[9 18] or have relied on antibody tests or immune

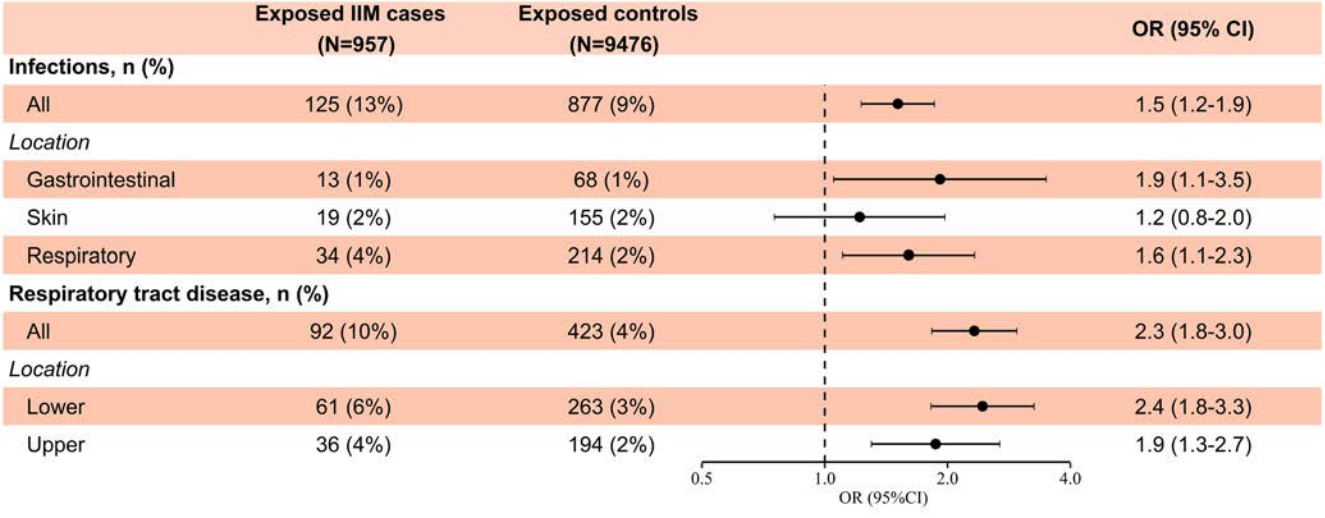


Figure 1 ORs for developing idiopathic inflammatory myopathies (IIM) associated with having a history of hospitalisation and/or outpatient visit for infections or respiratory tract disease (overall and by location). Groups are not mutually exclusive. Exposures occurring within 1 year before IIM diagnosis were not considered. ORs were estimated using conditional logistic regression conditioned on the matching set.

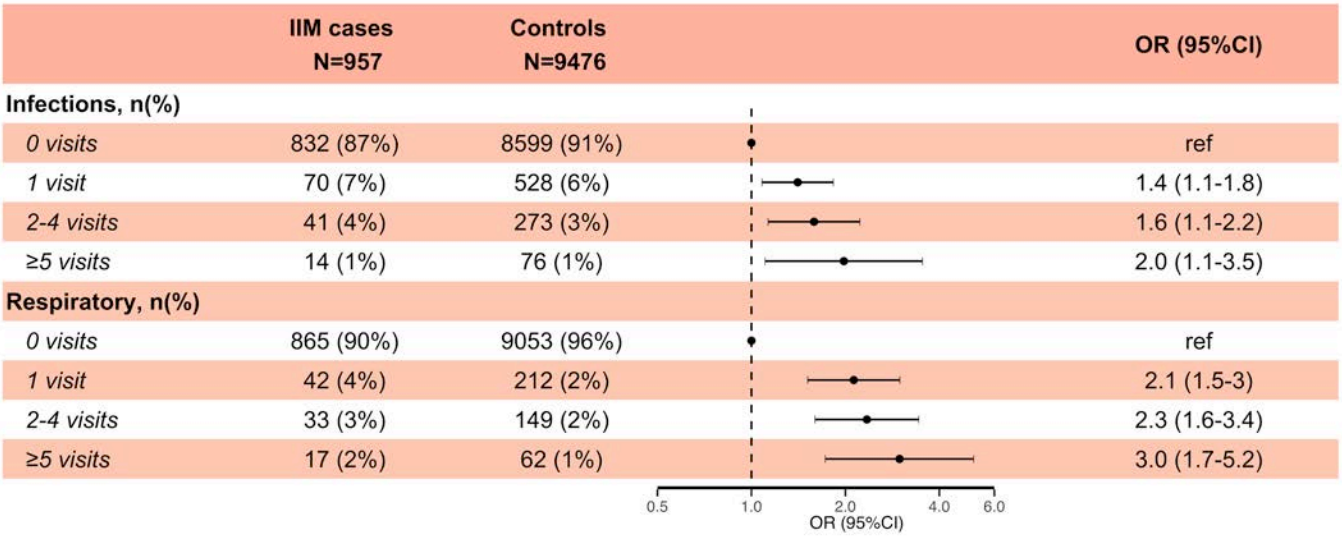


Figure 2 Effects of having multiple previous visits indicating infections or respiratory tract disease on the future risk to develop idiopathic inflammatory myopathies.CI, confidence interval; OR, odds ratios.

assays in prevalent cases,[19–21] making it difficult to infer anything about the temporality of the association. In a study with newly diagnosed PM and DM, similar findings as ours have been reported for infections (IRR=1.70).[8]

Respiratory inflammatory exposures were found to be associated with IIM in a recent study[15] but it only included three specific respiratory outcomes (sarcoidosis, pneumonia and tuberculosis) while our study includes a wider range of respiratory tract diseases. Their identification of IIM using hospitalisation data only likely led to the inclusion of more severe cases of IIM. Relying on questionnaire data may have caused some recall bias, and because the questionnaires were sent out years after diagnoses, only surviving patients could be included. Our findings suggest that respiratory tract disease is a stronger risk factor than infections. This could be due to the location of the inflammation and that the lung could act as the initial site of immune activation and auto-antibody production, as has been seen in RA.[12] In addition, the duration of the inflammation may play an important role as the increased inflammatory burden of chronic respiratory tract disease may lead to the break of self-tolerance.

This study has national coverage, enabling us to identify all individuals seeking care for IIM in Sweden and is to our knowledge the largest study to date investigating risk factors for IIM. In a recent study, we developed a stable algorithm to identify patients with IIM from the NPR.[16] The same study did, however, demonstrate the difficulties of discriminating between PM and IBM by only using ICD codes from the NPR. Therefore, only DM was analysed separately. If PM and IBM are affected differently by these exposures, we may have underestimated the effect on one and overestimated the effect on the other.

Using national register sources with long follow-up, it was possible to identify exposures occurring up to 10 years before diagnosis. However, with this design we did not identify triggers of disease occurring in close proximity to development of IIM but rather our aim was to identify exposures which occurred years before disease diagnosis. We believe an early time window is relevant for disease pathogenesis, as immunomodulation preceding development of clinically manifested disease could be initiated years before disease onset and therefore may be triggered by exposures occurring long before disease diagnosis. As has been demonstrated in other rheumatic diseases.[7 22] Furthermore, by excluding exposures occurring less than 1 year from IIM onset, we aimed to decrease the risk of reversed causality and detection bias. Also, we investigated if there was a difference

Table 2 Sensitivity analysis presenting number of identified idiopathic inflammatory myositis (IIM) cases and controls and corresponding ORs when varying latency time and excluding contributory diagnosis from the National Patient Register

	Exposed IIM cases (n=957)	Exposed controls (n=9476)	OR* (95% CI)
Infections, n (%)			
Main or contributory diagnosis, 1-year latency†	125 (13)	877 (9)	1.5 (1.2 to 1.9)
Main diagnosis, 1-year latency	104 (11)	724 (8)	1.5 (1.2 to 1.9)
Main or contributory diagnosis, 0-year latency	181 (19)	1081 (11)	1.9 (1.6 to 2.2)
Main or contributory diagnosis, 3-year latency	78 (11)	547 (7)	1.5 (1.2 to 1.9)
Respiratory tract disease, n (%)			
Main or contributory diagnosis, 1-year latency†	92 (10)	423 (4)	2.3 (1.8 to 3.0)
Main diagnosis, 1-year latency	68 (7)	334 (4)	2.1 (1.6 to 2.8)
Main or contributory diagnosis, 0-year latency	133 (14)	526 (6)	2.8 (2.3 to 3.4)
Main or contributory diagnosis, 3-year latency	56 (8)	256 (3)	2.3 (1.7 to 3.1)

*OR from conditional logistic regression.
†Definitions used in the main analyses. Only subjects with sufficient follow-up time included in the denominator when a latency period was included.

in the number of cases and controls who had no visits in the NPR prior to index date. Most individuals in the study had at least one previous visit in the NPR and even though there was a difference between cases and controls (13% vs 16% with no visits), we do not believe this explains the associations observed in our study.

Regarding the identification of exposures, the validity of the Swedish patient register is considered high with a positive predictive value (PPV) between 85%–95% for included diagnoses in one study.[23] For infections, both sepsis, pneumonia and infections of the central nervous system have a specificity over 95% while for respiratory tract diseases, asthma and COPD have a PPV of 75%–94% and 90%, respectively.[24 25] It would be interesting to further group exposures on specific type (eg, bacterial and viral) but as some ICD codes can be used for multiple types of infections, the specificity of these types would be too low.

Respiratory conditions could lead to a respiratory infection but because respiratory infections are only a small proportion of overall infections and results are similar between cases and controls, we believe this does not greatly influence our overall interpretation of the results (see online supplementary figure 1).

We conducted several sensitivity analyses to identify alternative explanations to these findings but the results remained the same. First, the associations found in this study may be due to an underlying cause of both IIM and infection or respiratory tract disease. Cases might have poorer health or an affected immune system increasing the risk of having the exposure. Therefore, previous healthcare consumption was adjusted for in the model as a proxy for general health. Also, individuals with IIM-associated conditions which could have a different disease aetiology were excluded. Second, because severity and timing of exposure could affect the associations, we varied the methods through which exposures were identified. When only hospitalisations were used to identify more severe exposures, it was possible to identify visits dating back to 1987. This did not affect the association for infections while the association with respiratory tract disease was weakened. Furthermore, because some infections and respiratory tract disease are treated in primary care, and may not have been captured, we likely missed some exposures. We addressed this by using drug dispensings, which enabled us to identify all exposures treated with prescription drugs. Still we could not identify exposures that were untreated, for example, common cold. We did observe a dose–response relationship between the number of visits and increased risk to develop IIM for both infections and respiratory tract disease which strengthens our hypothesis.

We cannot rule out that the observed associations are caused by residual confounding. Smoking has been suggested as a risk factor for a subgroup of IIM with anti-Jo-1 autoantibodies and is closely linked to respiratory conditions like COPD, asthma as well as upper respiratory conditions like chronic rhinitis. As we do not have smoking status available in these data, we were unable to adjust for this effect. However, we do not believe smoking would fully explain our results.

Our findings do not provide evidence of a protective effect of infections which has been suggested for autoimmune disease and allergies through the hygiene hypothesis.[26] Rather our findings imply that infections in the gastrointestinal and respiratory tracts can increase the risk of IIM. Molecular mimicry where a foreign antigen shares similarities with self-antigens, has been suggested as a potential mechanism.[5] Furthermore, gastrointestinal infections could change the gut microbiota, causing autoimmunity by affecting immunoregulatory mechanisms.[27] Infections also cause local inflammation as do many respiratory tract diseases like asthma, COPD and rhinitis and an increased inflammatory load or a general activation of the immune system could be involved in disease pathogenesis, possibly through priming of self-reactive lymphocytes and autoantibody production.[6 12]

In conclusion, we observed associations between IIM and infections of the gastrointestinal and respiratory tract as well as both upper and lower respiratory inflammatory conditions in our study. These findings suggest that external triggers of immunomodulation could be part of the aetiology of IIM and that these events could appear years before clinical manifestations of IIM.

Contributors Planning of the study: JS, EA, MH, IL. Analyses: JS. Interpretation of results: JS, EA, MH, IL. Writing: JS, EA, MH, IL.

Funding This study was supported by grants from the Swedish Society of Medicine, Swedish Research Council K2014-52X-14045-14-3, the Swedish Rheumatism Association, King Gustaf V 80 Year Foundation, Konung Gustaf V:s och Drottning Victorias Frimurarstiftelse, Funds at the Karolinska Institutet (KID) and through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet.

Competing interests IL has received honoraria from Bristol Myers Squibb and is currently receiving a research grant from Bristol Myers Squibb and from Astra Zeneca.

Ethics approval Ethics Committee at Karolinska Institutet.

REFERENCES

1 Dalakas MC. Polymyositis, dermatomyositis and inclusion-body myositis. *N Engl J Med* 1991;325:1487–98.
2 O'Hanlon TP, Carrick DM, Targoff IN, *et al*. Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1, and -DQA1 allelic profiles distinguish european american patients with different myositis autoantibodies. *Medicine* 2006;85:111–27.
3 Azali P, Barbasso Helmers S, Kockum I, *et al*. Low serum levels of vitamin D in idiopathic inflammatory myopathies. *Ann Rheum Dis* 2013;72:512–6.
4 Hengstman GJ, van Venrooij WJ, Vencovsky J, *et al*. The relative prevalence of dermatomyositis and polymyositis in Europe exhibits a latitudinal gradient. *Ann Rheum Dis* 2000;59:141–2.
5 Cusick MF, Libbey JE, Fujinami RS. Molecular mimicry as a mechanism of autoimmune disease. *Clin Rev Allergy Immunol* 2012;42:102–11.
6 Vanderlugt CL, Miller SD. Epitope spreading in immune-mediated diseases: implications for immunotherapy. *Nat Rev Immunol* 2002;2:85–95.
7 Carlens C, Jacobsson L, Brandt L, *et al*. Perinatal characteristics, early life infections and later risk of rheumatoid arthritis and juvenile idiopathic arthritis. *Ann Rheum Dis* 2009;68:1159–64.
8 Nielsen PR, Kragstrup TW, Deleuran BW, *et al*. Infections as risk factor for autoimmune diseases - A nationwide study. *J Autoimmun* 2016;74:176–81.
9 Rider LG, Wu L, Mamyrova G, *et al*. Environmental factors preceding illness onset differ in phenotypes of the juvenile idiopathic inflammatory myopathies. *Rheumatology* 2010;49:2381–90. keq277.
10 Koch MJ, Brody JA, Gillespie MM. Childhood polymyositis: a case-control study. *Am J Epidemiol* 1976;104:627–31.
11 Fathi M, Dastmalchi M, Rasmussen E, *et al*. Interstitial lung disease, a common manifestation of newly diagnosed polymyositis and dermatomyositis. *Ann Rheum Dis* 2004;63:297–301.
12 Mikuls TR, Payne JB, Deane KD, *et al*. Autoimmunity of the lung and oral mucosa in a multisystem inflammatory disease: the spark that lights the fire in rheumatoid arthritis? *J Allergy Clin Immunol* 2016;137:28–34.
13 Baka Z, Buzás E, Nagy G. Rheumatoid arthritis and smoking: putting the pieces together. *Arthritis Res Ther* 2009;11:238.
14 Chinoy H, Adimulam S, Marriage F, *et al*. Interaction of hla-drb1*03 and smoking for the development of anti-jo-1 antibodies in adult idiopathic inflammatory myopathies: a European-wide case study. *Ann Rheum Dis* 2012;71.
15 Helmers SB, Jiang X, Pettersson D, *et al*. Inflammatory lung disease a potential risk factor for onset of idiopathic inflammatory myopathies: results from a pilot study. *RMD Open* 2016;2:e000342.
16 Svensson J, Arkema EV, Lundberg IE, *et al*. Incidence and prevalence of idiopathic inflammatory myopathies in Sweden: a nationwide population-based study. *Rheumatology* 2017;56:802–10. In press.
17 Barlow L, Westergren K, Holmberg L, *et al*. The completeness of the Swedish Cancer register: a sample survey for year 1998. *Acta Oncol* 2009;48:27–33.
18 Pachman LM, Lipton R, Ramsey-Goldman R, *et al*. History of infection before the onset of juvenile dermatomyositis: results from the National Institute of

Arthritis and musculoskeletal and skin diseases Research Registry. *Arthritis Rheum* 2005;53:166–72.
19 Chen DY, Chen YM, Lan JL, *et al*. Polymyositis/dermatomyositis and nasopharyngeal carcinoma: the Epstein-Barr virus connection? *J Clin Virol* 2010;49:290–5.
20 Douche-Aourik F, Berlier W, Féasson L, *et al*. Detection of Enterovirus in human skeletal muscle from patients with chronic inflammatory muscle disease or Fibromyalgia and healthy subjects. *J Med Virol* 2003;71:540–7.
21 Gilbert DT, Morgan O, Smikle MF, *et al*. HTLV-1 associated polymyositis in Jamaica. *Acta Neurol Scand* 2001;104:101–4.
22 Rantapää-Dahlqvist S, de Jong BA, Berglin E, *et al*. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum* 2003;48:2741–9.
23 Ludvigsson JF, Andersson E, Ekbom A, *et al*. External review and validation of the swedish national inpatient register. *BMC Public Health* 2011;11:450.
24 Gedeborg R, Furebring M, Michaëlsson K. Diagnosis-dependent misclassification of infections using administrative data variably affected incidence and mortality estimates in ICU patients. *J Clin Epidemiol* 2007;60:155.e1–155.e11.
25 Örtqvist AK, Lundholm C, Wettermark B, *et al*. Validation of asthma and eczema in population-based swedish drug and patient registers. *Pharmacoepidemiol Drug Saf* 2013;22:850–60.
26 Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. *N Engl J Med* 2002;347:911–20.
27 Rook GA. Hygiene hypothesis and autoimmune diseases. *Clin Rev Allergy Immunol* 2012;42:5–15.

EXTENDED REPORT

Juvenile onset arthritis and pregnancy outcome: a population-based cohort study

Katarina Remaeus,[1,2] Kari Johansson,[1] Johan Askling,[1,3] Olof Stephansson[1,2]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210879).

[1]Department of Medicine Solna, Clinical Epidemiology Unit, Karolinska University Hospital and Institutet, Stockholm, Sweden
[2]Department of Women's and Children's Health, Division of Obstetrics and Gynaecology, Karolinska Institutet, Stockholm, Sweden
[3]Department of Rheumatology, Karolinska University Hospital, Stockholm, Sweden

Correspondence to
Katarina Remaeus, Department of Medicine, Solna, Clinical Epidemiology Unit, T2, Karolinska University Hospital and Institutet, 171 76 Stockholm, Sweden; katarina.remaeus@ki.se

Received 26 November 2016
Revised 27 April 2017
Accepted 3 June 2017
Published Online First 1 July 2017

ABSTRACT

Objectives Reports on pregnancy outcomes among women with juvenile onset arthritis (JIA) have been few and small. The aim of this study was to assess pregnancy outcomes in a large and contemporary cohort of women diagnosed with JIA.

Methods In a nationwide Swedish population-based cohort study between 1992 and 2011, we identified 1807 births among women with JIA and 1 949 202 control births. Since JIA is a heterogenic condition, births to women with JIA was divided into JIA paediatric only (n=1169) and JIA persisting into adulthood (n=638). ORs and 95% CIs were estimated with generalised estimating equations.

Results Women with JIA were at increased risk of preterm birth, especially medically indicated, in both subgroups: adjusted OR (aOR) 1.74 (1.35–2.67) for JIA paediatric and aOR 4.12 (2.76–6.15) for JIA persisting into adulthood. JIA persisting into adulthood was associated with very preterm birth (aOR 3.14, 1.58–6.24), spontaneous preterm birth (aOR 1.63, 1.11–2.39), small for gestational age birth (aOR 1.84, 1.19–2.85), early-onset pre-eclampsia (aOR 6.28, 2.68–13.81) and late-onset pre-eclampsia (aOR 1.96, 1.31–2.91). Women with JIA paediatric only were at increased risk of delivery by caesarean section (aOR 1.42, 1.66–1.73) and induction of labour (aOR 1.45, 1.18–1.77).

Conclusions We found increased risks of both maternal and infant complications among women with JIA confined to childhood and in women with JIA persistent into adulthood as compared with population controls. Pregnancies in women with JIA should thus be subject to increased surveillance during pregnancy and delivery.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) encompasses a heterogeneous group of clinical phenotypes characterised by onset of arthritis before the age of 16 years and is the most common childhood chronic rheumatic disease in the general population.[1] Although the natural course of JIA is variable and may include remission, more than one-third of all JIA persists into adulthood.[2]

Chronic inflammation such as rheumatoid arthritis and inflammatory bowel disease have been associated with increased risks for adverse pregnancy outcomes,[3–5] presumably linked to effects of ongoing and past inflammation. JIA may impact health in adulthood through several mechanisms: ongoing inflammatory activity and exposure to immune-modulatory therapies, systemic effects of past inflammatory activity such as impaired growth during adolescence and local effects such as joint destruction. So far, reports on pregnancy outcomes among women with JIA have been few: in 1991, Ostensen reported on 76 pregnancies in women with JIA and suggested an increased risk of caesarean delivery.[6] In 2013, Chen *et al* reported on 78 births in women with JIA and increased risks for caesarean delivery, pre-eclampsia and preterm birth.[7] Recently Ehrmann Feldman *et al* reported on a large cohort study of women with a history of JIA (n=1681). The study focused on neonatal outcomes and reported on higher proportions of prematurity, small for gestational age (SGA) and congenital malformations in the infants to women with a history of JIA.[8]

The aim of our study was to assess maternal and infant pregnancy outcomes in a large and contemporary cohort of women diagnosed with JIA and to the extent possible differentiate between JIA that did versus did not require medical attention at the time of pregnancy. To this end, and using national and population-based registers, we assessed pregnancy outcomes in women diagnosed with paediatric onset JIA and compared these with the expected pregnancy outcomes in the general population.

MATERIAL AND METHODS

We conducted a nationwide registry-based cohort study, including births in Sweden between 1992 and 2011.

Setting and data sources

Swedish healthcare is public and tax funded. There are national registries of prospectively collected information on health and social factors on all inhabitants, and a Civil Personal Registration (CPR) number (a unique 10-digit number assigned to each resident at birth or immigration) enables register linkages.[9] Paediatric rheumatology is managed by paediatricians with a subspecialisation or a special interest in rheumatology.

In this study, we used the following nationwide, population-based registers:

The Swedish Patient Register (PR) contains information on Swedish inpatient care since 1964 (nationwide since 1987).[10] On every visit, primary and contributory discharge diagnoses, as assigned by the discharging physician, are coded according to the calendar year—specific International Classification of Diseases (ICD) 7th through 10th revision codes, together with dates of admission and discharge. The completeness of this register is more than 99%. In 2001, nationwide data from specialist outpatient care were added to the PR. The same coding scheme is used as for the inpatient coding.

To cite: Remaeus K, Johansson K, Askling J, *et al*. *Ann Rheum Dis* 2017;**76**:1809–1814.

The Swedish Medical Birth Register (MBR) includes detailed information from more than 98% of all births in Sweden since 1973. Information is collected prospectively from the first visit during pregnancy to antenatal care, delivery and the neonatal period.[11] The register provides information on maternal, pregnancy and birth characteristics such as age, smoking, height, weight and parity, gestational age at birth, mode of delivery, delivery complications, stillbirth, birth weight, congenital malformations, Apgar score, and neonatal morbidity and mortality.

The Swedish Educational Register (ER) is held by Statistics Sweden and contains information on the highest level of completed education of all Swedish citizens 16–74 years of age. Information about foreign-born citizen's educational level is gathered yearly via surveys.

Identification of the study population

During the study period, 2 022 109 births were registered in the MBR. We excluded 59 259 births with multiple pregnancy and 2434 births without a CPR number. The remaining singleton births were merged with data from the PR and the ER. To avoid misclassification of the main exposure (diagnosis of juvenile onset arthritis), we excluded all births with a maternal diagnosis of connective tissue disease from the cohort (see online supplementary table S1 for ICD-codes used for exclusions). The final cohort consisted of 1 951 009 singleton births.

Maternal characteristics

Maternal characteristics are presented in table 1.

Exposure

In the PR, we identified all women (1090 unique women with 1807 births) with at a least one diagnosis of juvenile arthritis with onset before 18 years of age (see online supplementary table S2 for diagnosis codes). For simplicity hereafter entitled juvenile onset arthritis (JIA). We only accepted diagnoses assigned by physicians from internal medicine, rheumatology, paediatrics or paediatric surgery departments.

Our main interest was to explore the association between JIA and adverse pregnancy outcome. Since JIA is a heterogenic

Table 1 Descriptive data on 1 951 009 singleton births 1992–2011 with and without juvenile onset arthritis (JIA)

		Number of births		
		JIA n=1807	Population controls n=1 949 202	
Characteristic		n (%)	n (%)	p Value
Maternal age at delivery (years)				<0.0001
	13–24	489 (27.1)	327 085 (16.8)	
	25–29	614 (34.0)	641 705 (32.9)	
	30–34	502 (27.8)	636 029 (32.6)	
	35+	202 (11.2)	344 380 (17.7)	
	Missing	0 (0)	3 (<0.00)	
Calendar year of delivery				<0.0001
	1992–1998	470 (26.0)	692 609 (35.5)	
	1999–2004	465 (25.7)	534 360 (27.4)	
	2005–2011	872 (48.6)	722 233 (37.1)	
	Missing	0 (0)	0 (0)	
Mother's country of birth				<0.0001
	Nordic	1778 (98.4)	1 625 426 (83.7)	
	Non-Nordic	29 (1.6)	316 401 (16.3)	
	Missing	0 (0)	7375 (0.4)	
Parity				<0.0001
	Parous	871 (48.2)	1 107 165 (56.8)	
	Primiparous	936 (51.8)	842 037 (43.2)	
	Missing	0 (0)	0 (0)	
BMI (kg/m^2)				<0.0001
	11.0–18.4	77 (4.9)	43 440 (2.6)	
	18.5–24.9	1006 (63.8)	1 067 310 (63.2)	
	25.0–29.9	335 (21.2)	408 432 (24.2)	
	30.0+	159 (10.1)	169 721 (10.1)	
	Missing	230 (12.7)	260 299 (13.4)	
Smoking habits (number of cigarettes per day)				0.0014
	Non-smokers	1457 (85.6)	1 628 704 (88.1)	
	1–9	179 (10.5)	150 576 (8.1)	
	10+	67 (3.9)	69 466 (3.8)	
	Missing	104 (5.8)	100 456 (5.2)	
Maternal education				0.0008
	>High school	752 (42.4)	877 171 (46.4)	
	≤High school	1021 (57.6)	1 014 582 (53.6)	
	Missing	34 (1.9)	57 449 (2.9)	

BMI, body mass index.

condition that may or may not persist into adulthood, we constructed two subgroups of births within the JIA-group.

The first subgroup, 'JIA paediatric only' (considered to have JIA confined to childhood and adolescence) included births to women with onset of JIA before 18 years of age or a diagnosis of any of the corresponding adult arthritis diagnoses: psoriatic arthritis, ankylosing spondylitis, inflammatory spondyloarthropathies or rheumatoid arthritis before the age of 18 years (the Swedish cut-off between paediatric and adult care), but without any visits or hospitalisations listing either of these diagnoses after the age of 18 years until the time point of delivery (n=1169).

The second subgroup 'JIA persisting into adulthood' included births to women with a diagnosis of JIA, with the same criteria as described above and who also had at least one visit or hospitalisation listing for either of these diagnoses after the age of 18 years but before delivery (n=638) (see online supplementary table S3 for numbers of JIA in each group and diagnosis of arthritis >18 years).

Outcomes

Pre-eclampsia (including eclampsia) was analysed among all births (including stillbirths) and identified by the ICD-9 codes 642E-G and ICD-10 codes O14 and O15. We used gestational age at delivery to categorise pre-eclampsia as early onset (before gestational week 34+0) and late onset (gestational age ≥34+0). Among all births, we assessed stillbirth (intrauterine death after gestational week 28 from 1992 to June 30 2008 and from July 1 from gestational week 22).

All other outcomes were analysed among live births. Preterm birth (before 37 gestational weeks) was subcategorised into very preterm (<32+0 gestational weeks) and moderately preterm (32+0–36+6 gestational weeks) and by type of onset: spontaneous or induced. Onset of labour was categorised into spontaneous or induced labour and mode of delivery into vaginal birth, assisted vaginal birth or caesarean section (emergency or planned).

For fetal outcomes, we assessed neonatal death (death before 28 days), Apgar score <7 at 5 min and SGA. SGA was defined as a birth weight of more than 2 SD below the sex-specific mean for gestational age, based on the Swedish reference curve of estimated fetal growth.[12]

Statistical analyses

Differences between births in mothers with JIA and the reference population with regards to categorical data were assessed by a χ^2 test where $p<0.05$ was considered statistically significant. Unconditional logistic regression analyses were performed to estimate crude and adjusted ORs (aORs) with associated 95% CIs for outcome variables. The generalised estimation equation method was used in the model to account for the clustering due to the inclusion of multiple births from the same mother. In the adjusted analyses, we controlled for maternal age at delivery, parity, body mass index (BMI), smoking habits, educational level, the mother's country of birth and calendar year of birth.

We used a formal interaction test in the model to estimate possible effect modification between the exposure diabetes and the outcome pre-eclampsia. We performed sensitivity analyses where we defined JIA as at least two JIA diagnoses before 18 years of age (online supplementary table S4), restricted the analyses to births during 2002–2011 (online supplementary table S5) and restricted the identification of JIA diagnoses to 1987–2011 (ICD-9 and ICD-10 codes) (online supplementary table S6). We also performed a sensitivity analysis regarding time since last diagnosis in women with JIA paediatric restricted to year 2003 and onwards, with requirement of 2 years from last diagnosis until delivery. All analyses were conducted using SAS software, V.9.4.

RESULTS

During the study period, there were 1807 singleton births in women with a diagnosis of JIA before delivery and 1 949 202 population control births. Compared with population controls, women with JIA were generally younger, leaner, more often primiparous and born in the Nordic countries. They were more often smokers and also had a lower education level (table 1). Mean and median age at delivery was 28 years in both JIA subgroups. Mean time from last visit until delivery in women with JIA paediatric only after 2003 (n=563) was 18.4 years (SD 9.1) and median time was 20.1 years (range 10–26). In women with JIA persisting into adulthood, mean time from first diagnosis until delivery was 17.7 years (SD 7.8) and median was 18.0 years (range 11–23).

JIA paediatric only

Preterm birth was more common in the JIA paediatric group (5.9%) than in population controls (5.0%). There was an elevated risk of moderately preterm birth (aOR 1.43; 95% CI 1.07 to 1.91) and medically indicated preterm birth (aOR 1.74; 95% CI 1.35 to 2.67) in women with JIA paediatric only compared with population controls. Regardless of gestational age, a larger proportion of the JIA paediatric only births were induced: 13.9% compared with 10.3% in the population control group (aOR 1.45; 95% CI 1.18 to 1.77). Caesarean delivery was more common in women with JIA paediatric only (aOR 1.42; 95% CI 1.66 to 1.73), with increased risks of emergency and elective caesarean sections (table 2). There was no difference in risk of pre-eclampsia, assisted vaginal delivery, stillbirth, 5 min Apgar score, SGA birth and neonatal mortality between the groups.

JIA persisting into adulthood

Compared with population controls, there was an increased risk of pre-eclampsia in women with JIA persisting into adulthood (table 3). When stratifying by time of onset, we found increased risks for early onset pre-eclampsia (aOR 6.28; 95% CI 2.86 to 13.81) and late onset pre-eclampsia (aOR 1.96; 95% CI 1.31 to 2.91) in the JIA persisting group compared with population controls (table 3). The risk of preterm birth was increased (aOR 2.40; 95% CI 1.81 to 3.18) and also when divided into moderately preterm and very preterm birth (aOR 2.74; 95% CI 1.68 to 3.08 and aOR 3.14; 95% CI 1.58 to 6.24, respectively) (table 3). Both the risk of medically indicated and spontaneous onset of preterm birth was increased in women with JIA persisting into adulthood, and the births in this group were also induced to a greater extent compared with population control births (aOR 1.37; 95% CI 1.07 to 1.75). In addition, they were more often delivered by caesarean section, whereas there was no difference regarding assisted vaginal delivery (table 3). The proportion of pre-eclampsia in preterm birth was 22.9% in the JIA persisting group compared with 12.4% in population control births. In medically indicated preterm birth, the proportion of pre-eclampsia was 45.5% compared with 36.6% in population control births and in SGA births 32.1% compared with 14.3% (table 4).

There was no difference in risk regarding stillbirth in JIA persisting into adulthood compared with population controls, whereas there was an increased risk of SGA (aOR 1.84; 95% CI 1.19 to 2.50) (table 3). The other neonatal outcomes did not differ between the groups.

Table 2 Adverse pregnancy and infant outcomes among singleton births from 1992 to 2011 with JIA paediatric only, compared with population control births

Outcome variable among all births	JIA paediatric only n=1169	Population control births n=1 949 840	OR	(95% CI)	aOR*	(95% CI)
	N (%)	N (%)				
Pre-eclampsia	40 (3.4)	54 198 (2.8)	1.24	0.89 to 1.72	1.27	0.90 to 1.80
Early onset pre-eclampsia†	4 (0.3)	4870 (0.3)	1.37	0.51 to 3.67	1.76	0.66 to 4.68
Late onset pre-eclampsia‡	36 (3.1)	49 241 (2.6)	1.22	0.87 to 1.71	1.22	0.85 to 1.75
Stillbirth	5 (0.4)	6443 (0.3)	1.30	0.54 to 3.10	1.44	0.54 to 3.81
Outcome variables among live births	**JIA paediatric only n=1164**	**Population control live births n=1 943 397**	**OR**	**(95% CI)**	**aOR***	**(95% CI)**
	N (%)	N (%)				
Preterm delivery (<37+0)	69 (5.9)	96 423 (5.0)	1.21	0.93 to 1.57	1.32	1.00 to 1.76
Moderately preterm delivery	66 (5.7)	83 174 (4.3)	1.34	1.03 to 1.75	1.43	1.07 to 1.91
Very preterm delivery	3 (0.3)	13 249 (0.7)	0.38	0.12 to 1.17	0.52	0.16 to 1.62
Preterm delivery, medically indicated	26 (2.2)	29 046 (1.5)	1.51	1.01 to 2.25	1.74	1.35 to 2.67
Preterm delivery, spontaneous	43 (3.7)	67 377 (3.5)	1.07	0.76 to 1.50	1.14	0.79 to 1.64
Induction of labour	158 (13.9)	196 383 (10.3)	1.40	1.17 to 1.68	1.45	1.18 to 1.77
Assisted vaginal delivery	88 (9.1)	141 630 (8.5)	1.08	0.86 to 1.35	1.33	0.87 to 1.46
Caesarean section	198 (17.0)	279 984 (14.4)	1.22	1.02 to 1.46	1.42	1.66 to 1.73
Emergency caesarean section	104 (8.9)	145 651 (7.5)	1.21	0.98 to 1.50	1.37	1.08 to 1.74
Elective caesarean section	94 (8.1)	134 333 (6.9)	1.18	0.93 to 1.50	1.39	1.08 to 1.78
Apgar <7 at 5 min	8 (0.7)	20 356 (1.1)	0.65	0.33 to 1.31	0.71	0.34 to 1.50
Small for gestational age	26 (2.2)	50 687 (2.6)	0.85	0.58 to 1.25	0.86	0.56 to 1.32
Neonatal death (0–27 days)	1 (0.1)	3600 (0.2)	0.46	0.07 to 3.29	0.67	0.09 to 4.73

*Adjusted for maternal age at delivery, parity, BMI, calendar year of birth, smoking habits, educational level and the mother's country of birth.
†Early onset before gestational week 34+0.
‡Late onset from gestational week 34+0.
aOR, adjusted; OR; BMI, body mass index.

Since diabetes mellitus is a known risk factor of pre-eclampsia, mostly late-onset pre-eclampsia,[13] but is also associated with juvenile arthritis,[14] we analysed diabetes both as a confounder and as an interaction term in order to rule out effect modification. However, none of the analyses changed the point estimates from those of the original analysis.

In sensitivity analyses restricted to individuals with at least two JIA diagnoses codes, to birth 2002–2011 or to JIA identified 1987–2011, the relative risks associated with JIA in the paediatric only and persistent groups did not differ substantially from the original model (see online supplementary tables S4–S6).

DISCUSSION

Main findings

In this study, we found increased risks of both maternal and neonatal complications among women with JIA compared with population controls. The increased risks were seen in both JIA subgroups. Generally, however, the associations were more pronounced in the JIA persisting into adulthood group. We did not find any association between JIA (regardless of persistence into adulthood) and neonatal death, stillbirth or low Apgar score.

Interpretation

There was a strong association between JIA persisting into adulthood and pre-eclampsia. Furthermore, pre-eclampsia was correlated to the increased risks of caesarean section, medically indicated preterm birth and SGA, in women with persisting JIA. Pre-eclampsia is characterised by abnormal placentation with subsequent maternal inflammatory and vascular response, especially early onset pre-eclampsia, and is often accompanied by intrauterine growth restriction.[15] Abnormal placentation can result in pregnancy complications, such as miscarriage, preterm birth, pre-eclampsia and SGA birth.[16 17] Our results, showing a near sixfold increase in early onset pre-eclampsia, increased risks of both moderately and very preterm birth with spontaneous and medically indicated onset and infants born SGA in women with JIA persisting into adulthood are in line with a model presented by Ostensen *et al* regarding women with rheumatic disease and high-risk pregnancies.[17] Also prolonged treatment with corticosteroids has been associated with multiple adverse pregnancy outcomes including risk of preterm delivery and intrauterine growth restriction.[18] In a prospective cohort of women with RA, de Man *et al* concluded that pregnancy outcome in women with well-controlled RA is comparable with that in the general population.[19] The effect of prednisone on birth weight was mediated by a lower gestational age at delivery, whereas a higher disease activity independently influenced birth weight negatively, suggesting an immune-mediated mechanism. The adverse outcomes seen in our study, more pronounced in JIA persisting, may be due to several factors including disease activity or medication during pregnancy.

We have analysed the proportion of women with preterm birth, medically indicated preterm birth and SGA births because of pre-eclampsia. We found that the proportions of pre-eclampsia in JIA paediatric only subgroup were comparable with those in the population control births. In opposite, proportions of pre-eclampsia in the JIA persisting into adulthood subgroup were generally higher than in population control

Table 3 Adverse pregnancy and infant outcomes among singleton births from 1992 to 2011, with JIA persisting into adulthood, compared with population control births

Outcome variable among all births	JIA persisting n=638	Population control births n=1 950 371				
	N (%)	N (%)	OR	95 % CI	aOR*	95 % CI
Pre-eclampsia	41 (6.4)	54 197 (2.8)	2.40	1.72 to 3.35	2.31	1.61 to 3.32
Early onset preeclampsia†	9 (1.4)	4865 (0.3)	5.72	2.76 to 11.85	6.28	2.86 to 13.81
Late onset preeclampsia‡	32 (5.2)	49 245 (2.6)	2.08	1.44 to 3.01	1.96	1.31 to 2.91
Stillbirth	3 (0.5)	6445 (0.3)	1.42	0.46 to 4.42	1.33	0.33 to 5.32
Outcome variables among live births	**JIA persisting n=635**	**Population control live births n=1 943926**				
	N (%)	N (%)	OR	95% CI	aOR*	95% CI
Preterm delivery (<37+0 gw)	70 (11.2)	96 422 (5.0)	2.37	1.83 to 3.08	2.40	1.81 to 3.18
Moderately pre-term delivery	58 (9.3)	83 182 (4.3)	2.28	1.72 to 3.02	2.27	1.68 to 3.08
Very preterm delivery	12 (1.9)	13 240 (0.7)	2.81	1.52 to 5.20	3.14	1.58 to 6.24
Preterm delivery, medically indicated	33 (5.2)	29 039 (1.5)	3.61	2.50 to 5.23	4.12	2.76 to 6.15
Preterm delivery, spontaneous	37 (5.8)	67 383 (3.5)	1.72	1.22 to 2.43	1.63	1.11 to 2.39
Induction of labour	90 (14.4)	196 451 (10.3)	1.46	1.16 to 1.84	1.37	1.07 to 1.75
Assisted vaginal delivery	41 (9.1)	141 677 (8.5)	1.08	0.77 to 1.51	0.97	0.68 to 1.38
Caesarean section	186 (29.3)	279 996 (14.4)	2.46	2.00 to 3.03	2.47	1.99 to 3.08
Elective caesarean section	113 (17.8)	134 314 (6.9)	2.92	2.29 to 3.71	3.01	2.32 to 3.90
Emergency caesarean section	73 (11.5)	145 682 (7.5)	1.60	1.25 to 2.06	1.57	1.19 to 2.08
Apgar <7 at 5 min.	10 (1.6)	20 354 (1.1)	1.51	0.81 to 2.81	1.25	0.62 to 2.50
Small for gestational age	28 (4.4)	50 685 (2.6)	1.72	1.15 to 2.58	1.84	1.19 to 2.85
Neonatal death (0–27 days)	0 (0.0)	3600 (0.2)	NA	NA	NA	NA

*Adjusted for maternal age at delivery, parity, BMI, calendar year of birth, smoking habits, educational level and mother's country of birth.
†Early onset before gestational week 34+0.
‡Late onset from gestational week 34+0.
aOR, adjusted OR; BMI, body mass index; JIA, juvenile idiopathic arthritis.

births indicating that management due to pre-eclampsia influences other adverse outcomes as well. A percentage of 45.5 of medically indicated preterm births also had pre-eclampsia, and this is a major contributor to the increased rate of induction of labour or caesarean section in women with JIA persisting into adulthood. There was also a difference compared with control births in the analysis of SGA births where 32.1% had pre-eclampsia and only 14.3% of the control births.

The increased risk of caesarean delivery is in line with findings from previous studies.[6 7 20] In our study, we confirm this finding in both subgroups of JIA. The associations were stronger for elective than emergency caesarean sections. Ostensen has earlier

Table 4 Proportions of pre-eclampsia by adverse outcomes for live births in women with JIA paediatric only compared with population controls and JIA persisting into adulthood compared with population controls

	JIA paediatric only		JIA persisting into adulthood	
	No (n (%))	Yes (n (%))	No (n (%))	Yes (n (%))
Pre-eclampsia	Preterm birth (n=96 492)			
No (n (%))	84 500 (87.6)	61 (88.4)	84 507 (87.6)	54 (77.1)
Yes (n (%))	11 923 (12.4)	8 (11.6)	11 915 (12.4)	16 (22.9)
	Medically indicated preterm birth (n=29 072)			
No (n (%))	18 415 (63.4)	18 (69.2)	18 415 (63.4)	18 (54.6)
Yes (n (%))	10 631 (36.6)	8 (30.8)	10 624 (36.6)	15 (45.5)
	Caesarean delivery (n=280 182)			
No (n (%))	259 909 (92.8)	181 (91.4)	259 925 (92.8)	165 (88.7)
Yes (n (%))	20 075 (7.2)	17 (8.6)	20 071 (7.2)	21 (11.3)
	Elective caesarean delivery (n=134 427)			
No (n (%))	123 752 (92.1)	87 (92.6)	123 741 (92.1)	98 (86.7)
Yes (n (%))	10 581 (7.9)	7 (7.5)	10 573 (7.9)	15 (13.3)
	SGA birth (n=50 713)			
No (n (%))	43 413 (85.7)	23 (88.5)	43 417 (85.7)	19 (67.9)
Yes (n (%))	7274 (14.4)	3 (11.5)	7268 (14.3)	9 (32.1)

The reference groups (no diagnosis) include all women without JIA paediatric only and JIA persisting into adulthood, respectively.
JIA, juvenile onset arthritis; SGA, small for gestational age.

described increased rates of caesarean sections in patients with JIA; in 15 of 20 caesarean deliveries, the indication was related to the juvenile arthritis disease.[6]

Chen *et al* did not find an increased risk of induction of labour in their study,[7] which is in contrast to our results. It may reflect different obstetric routines. In Chen's study, 21.8% of pregnancies with JIA were induced, which did not significantly differ from the reference population. In our cohort, approximately 14% of the JIA pregnancies were induced compared with 10% in population controls. The same reasoning can be applied to assisted vaginal delivery, where 17% of pregnancies in the Chen study underwent assisted vaginal delivery compared with 8% and 9% in the subgroups herein, respectively.

In the present study, women with JIA persisting into adulthood exhibited an increased risk of spontaneous preterm birth whereas this was not observed in JIA paediatric only group. In the study by Chen, 13.4% of women with juvenile onset arthritis delivered spontaneously preterm, compared with 5% in the reference population. We also found increased risks of medically indicated preterm birth. The recent study by Feldman *et al* reported on adverse neonatal outcomes, which included both preterm birth, SGA and congenital malformations. The proportion of preterm birth were 9.2% in their JIA group and 7.5% in the non-JIA-group, but they were not analysed by spontaneous or induced onset. The finding of an association between JIA and major congenital malformation is new and has not been addressed in our study.

Strengths and limitations of the study

The large study sample made it possible to stratify the exposure of JIA based on persistence into adulthood or not. We were able to study early and late pre-eclampsia onset, which provides new and important information. The population-based design generates a high generalisability for women with JIA. We included women with a diagnosis prior to delivery and also only accepted diagnoses noted by specialists. We excluded women with JIA who later received diagnosis codes indicating connective tissue disease that is known to be strongly associated with adverse pregnancy outcome. Information on JIA was collected in the PR, ruling out recall bias. Furthermore, we were able to analyse a large number of confounders prospectively collected during pregnancy.

There are several limitations to our study. Most importantly we did not have information of medication before or during pregnancy or disease activity why we cannot fully ensure remission in JIA paediatric group. Results from a sensitivity analysis of time from last visit to delivery was more than 18 years after 2003, which is reassuring. Also, we were not able to analyse data according to the JIA subtypes as defined, for example, by ILAR. There is no specific validation of JIA diagnosis in Sweden, validation studies of, for example, adult rheumatoid arthritis suggest a 90% positive predictive value of the corresponding diagnosis codes.[21]

Conclusion

In conclusion, JIA is associated with adverse pregnancy outcomes including preterm birth, pre-eclampsia, increased rates of induction of labour and caesarean section, in particular in women whose JIA is persistent into adulthood. The absolute risk of pre-eclampsia in women with JIA increase from 2.8% to 3.4% in paediatric JIA and to 6.4% in women with persisting JIA. Women with JIA should be subject to increased surveillance during pregnancy and delivery.

Contributors Study concept and design: KR, OS and JA. Acquisition of data: OS. Statistical analyses: KR and KJ. Drafting of the manuscript: KR and OS. Critical revision of the manuscript for important intellectual content: KR, KJ, JA and OS.

Funding JA was funded by the Swedish Cancer Society (Cancerfonden), the Swedish Foundation for Strategic Research (SSF), the Stockholm County Council (ALF), the Swedish Heart-Lung Foundation, the Swedish Research Council (Vetenskapsrådet), and the Nordic Research Council (Nordforsk). JA has research ongoing agreements with Lilly, Pfizer, Samsung, UCB, Roche, Merck. OS was supported by grants from the Swedish Research Council (grant 2013-2429), the Stockholm County Council and the Strategic Research Program in Epidemiology at Karolinska Institutet.

Competing interests KR, KJ and OS none declared. JA reports grants from Astra Zeneca, grants from Pfizer, grants from Abbvie, grants from MSD, grants from Roche, grants from UCB, grants from Lilly, grants from Samsung, outside the submitted work and pertaining to safety surveillance of anti-rheumatic drugs using Swedish registers.

Patient consent The study is register based, not based on individuals.

Ethics approval Ethical approval from Karolinska Institutet, Ethics committee Dnr 2006889-31, Dnr 2007/1391-32,Dnr 2008/631-32,Dnr 2009/1853-32

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011;377:2138–49.
2 Coulson EJ, Ng WF, Goff I, *et al*. Cardiovascular risk in juvenile idiopathic arthritis. *Rheumatology* 2013;52:1163–71.
3 de Man YA, Dolhain RJ, Hazes JM. Disease activity or remission of rheumatoid arthritis before, during and following pregnancy. *Curr Opin Rheumatol* 2014;26:329–33.
4 Bröms G, Granath F, Linder M, *et al*. Complications from inflammatory bowel disease during pregnancy and delivery. *Clin Gastroenterol Hepatol* 2012;10:1246–52.
5 Bröms G, Granath F, Linder M, *et al*. Birth outcomes in women with inflammatory bowel disease: effects of disease activity and drug exposure. *Inflamm Bowel Dis* 2014;20:1091–8.
6 Ostensen M. Pregnancy in patients with a history of juvenile rheumatoid arthritis. *Arthritis Rheum* 1991;34:881–7.
7 Chen JS, Ford JB, Roberts CL, *et al*. Pregnancy outcomes in women with juvenile idiopathic arthritis: a population-based study. *Rheumatology* 2013;52:1119–25.
8 Ehrmann Feldman D, Vinet É, Bernatsky S, *et al*. Birth outcomes in women with a history of Juvenile Idiopathic Arthritis. *J Rheumatol* 2016;43:804–9.
9 Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, *et al*. The swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol* 2009;24:659–67.
10 Ludvigsson JF, Andersson E, Ekbom A, *et al*. External review and validation of the swedish national inpatient register. *BMC Public Health* 2011;11:450.
11 The Swedish Centre for Epidemiology. The Swedish Medical Birth Register: a summary of content and quality. 2003 http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/10655/2003-112-3_20031123.pdf (cited 2015 1st Aug).
12 Marsál K, Persson PH, Larsen T, *et al*. Intrauterine growth curves based on ultrasonically estimated foetal weights. *Acta Paediatr* 1996;85:843–8.
13 Lisonkova S, Joseph KS. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. *Am J Obstet Gynecol* 2013;209:544.e1–544.e12.
14 Hermann G, Thon A, Mönkemöller K, *et al*. Comorbidity of type 1 diabetes and juvenile idiopathic arthritis. *J Pediatr* 2015;166:e933:930–5.
15 Redman CW, Sargent IL. Immunology of pre-eclampsia. *Am J Reprod Immunol* 2010;63:534–43.
16 Young BC, Levine RJ, Karumanchi SA. Pathogenesis of preeclampsia. *Annu Rev Pathol* 2010;5:173–92.
17 Østensen M, Andreoli L, Brucato A, *et al*. State of the art: reproduction and pregnancy in rheumatic diseases. *Autoimmun Rev* 2015;14:376–86.
18 Østensen M, Förger F. How safe are anti-rheumatic drugs during pregnancy? *Curr Opin Pharmacol* 2013;13:470–5.
19 de Man YA, Hazes JM, van der Heide H, *et al*. Association of higher rheumatoid arthritis disease activity during pregnancy with lower birth weight: results of a national prospective study. *Arthritis Rheum* 2009;60:3196–206.
20 Packham JC, Hall MA. Long-term follow-up of 246 adults with juvenile idiopathic arthritis: social function, relationships and sexual activity. *Rheumatology* 2002;41:1440–3.
21 Waldenlind K, Eriksson JK, Grewin B, *et al*. Validation of the rheumatoid arthritis diagnosis in the Swedish National Patient Register: a cohort study from Stockholm County. *BMC Musculoskelet Disord* 2014;15:432.

EXTENDED REPORT

Persons with chronic widespread pain experience excess mortality: longitudinal results from UK Biobank and meta-analysis

Gary J Macfarlane,[1,2] Maxwell S Barnish,[1,2] Gareth T Jones[1,2]

[1]Epidemiology Group, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK
[2]Aberdeen Centre for Arthritis and Musculoskeletal Health University of Aberdeen, Aberdeen, UK

Correspondence to
Professor Gary J Macfarlane, Health Sciences Building Foresterhill, Aberdeen,United KingdomAB25 2ZD; g.j.macfarlane@abdn.ac.uk

Received 15 March 2017
Revised 5 May 2017
Accepted 3 June 2017
Published Online First 20 July 2017

ABSTRACT

Objective It is uncertain whether persons with chronic widespread pain (CWP) experience premature mortality. Using the largest study conducted, we determine whether such a relationship exists, estimate its magnitude and establish what factors mediate any relationship.

Methods UK Biobank, a cohort study of 0.5 million people aged 40–69 years, recruited throughout Great Britain in 2006–2010. Participants reporting 'pain all over the body' for >3 months were compared with persons without chronic pain. Information on death (with cause) was available until mid-2015. We incorporated these results in a meta-analysis with other published reports to calculate a pooled estimate of excess risk.

Results 7130 participants reported CWP and they experienced excess mortality (mortality risk ratio 2.43, 95%CI 2.17 to 2.72). Specific causes of death in excess were cancer ($1.73_{\text{adjusted age and sex}}$, 95% CI 1.46 to 2.05), cardiovascular ($3.24_{\text{adjusted age and sex}}$, 95% CI 2.55 to 4.11), respiratory ($5.66_{\text{adjusted age and sex}}$, 95% CI 4.00 to 8.03) and other disease-related causes ($4.04_{\text{adjusted age and sex}}$, 95% CI 3.05 to 5.34). Excess risk was substantially reduced after adjustment for low levels of physical activity, high body mass index, poor quality diet and smoking. In meta-analysis, all studies showed significant excess all-cause (combined estimate 1.59 (95% CI 1.05 to 2.42)), cardiovascular and cancer mortality.

Conclusions Evidence is now clear that persons with CWP experience excess mortality. UK Biobank results considerably reduce uncertainty around the magnitude of excess risk and are consistent with the excess being explained by adverse lifestyle factors, which could be targeted in the management of such patients.

INTRODUCTION

Persons with CWP, the characteristic symptom of fibromyalgia (FM), have been reported to experience premature mortality. The original observation, in a UK study, found 30% excess mortality was explained primarily by increased cancer incidence and reduced survival.[1 2] A subsequent UK study confirmed the 30% excess mortality was primarily from increased cancer and cardiovascular deaths.[3]

Studies to identify the mediators of such a relationship have focused on low levels of physical activity, since the specific cancers contributing to excess mortality (female breast, prostate and colon) have been linked to low physical activity.[4 5] It has been hypothesised that CWP may lead to low levels of physical activity and this was confirmed by a longitudinal study.[6] Further studies have suggested additional lifestyle mediators of excess mortality: overweight has been shown to predict CWP onset and persistence[7 8]; persons with CWP have been reported as more likely to smoke, and women with CWP have been shown to have poorer quality diet.[9]

However not all studies conducted have found an excess mortality among persons with CWP. Meta-analyses have reported considerable heterogeneity, which has been attributed to differences in study populations, follow-up time, pain phenotype, methods of analysis and use of confounding factors.[10 11] Currently there is considerable uncertainty as to whether there is an excess mortality risk. It is important to determine whether an excess risk exists and if so to quantify it, since there remains the potential, as part of managing patients with CWP or FM, to modify the mediators of any excess risk.

We therefore now report on the largest study to examine the relationship between chronic widespread pain (CWP) and mortality experience, and with considerably more detailed information on potential mediators of any excess risk. Further we include these results in a meta-analysis, with other published reports, to evaluate the coherence of evidence.

METHODS

UK Biobank

Detailed methods used by UK Biobank have been published previously,[12] and we provide only summary details of relevance to the current analysis. The study aimed to recruit around half a million persons aged 40–69 years who were registered with a general practitioner within the National Health Service. Approximately 9.2 million invitations were issued, between 2006 and 2010, to people living within 25 miles of one of 22 assessment centres throughout Great Britain.

At the assessment centre, participants completed questionnaires including items on lifestyle and environment. Information on pain was collected by means of a touchscreen questionnaire. Participants were asked 'In the last month have you experienced any of the following that interfered with your usual activities?' If they answered positively, they were then provided with a list that included individual regional pain sites, or alternatively they could choose the response 'pain all over the body'. Subjects who reported 'pain all over the body' were not offered the option of choosing any further

To cite: Macfarlane GJ, Barnish MS, Jones GT. *Ann Rheum Dis* 2017;**76**:1815–1822.

regional sites. Respondents were asked whether the reported pain had lasted at least 3 months, and those with 'pain all over the body' which had lasted 3 months were defined as having CWP. Participants were identified on the Office for National Statistics (ONS) records. ONS collects information on cause of death from civil registration records. For registered deaths, the underlying cause of death is derived from the sequence of conditions leading directly to the death and is recorded on the death certificate. The current analysis uses data on vital status up to August 2015.

The exposures that we considered in terms of mediating any relationship between CWP and mortality were focused on factors potentially modifiable as part of the management of CWP:

- Body mass index (BMI), derived from measured height and weight, categorised according to standard cut-offs of the WHO.
- Physical activity: minutes of walking per week ('In a typical week, on how many days did you walk for at least 10 mins at a time' and 'How many minutes did you usually spend walking on a typical day?'); minutes of moderate activity per week ('In a typical week, on how many days did you do 10 mins or more of moderate physical activities like carrying light loads, cycling at a normal pace (do not include walking)' and 'How many minutes did you usually spend doing moderate activities on a typical day?'); minutes of vigorous activity per week (as before but vigorous defined as 'activities that make you sweat or breathe had such as fast cycling, aerobics, heavy lifting'). These were categorised as nil and then by quartiles.
- Diet: participants were asked: (1) 'On average how many heaped tablespoons of cooked vegetables would you eat per day? (Do not include potatoes.)' (2) 'On average how many heaped tablespoons of salad or raw vegetables would you eat per day? (Include lettuce and tomato in sandwiches)' (3) 'About how many pieces of fresh fruit would you eat per day?' (4) 'About how many pieces of dried fruit would you eat per day?' Total daily 'portions' of cooked vegetables, raw vegetables and salad consumption were calculated and recoded as quintiles. Frequency of alcohol consumption was determined with response categories: never, daily or almost daily, three or four times a week, once or twice a week, one to three times a month, special occasions only. The latter two categories were combined into 'Less frequently than once or twice per week'.
- Smoking status: a history of smoking was recorded, which allowed us to classify respondents as current, never (or very rare) or ex-smokers, the latter group being divided into ex-regular and ex-occasional smokers.

UK Biobank analysis

We used Poisson regression models with robust estimation of SEs to model the relationship between CWP and all-cause mortality, adjusted for age group and sex. We tested and confirmed that the mediating variables were not collinear. We compared persons with CWP to persons who did not report any chronic pain. We additionally examined specific major causes of death as outcomes, including cardiovascular disease, respiratory disease and cancer. We report the mortality risk ratio (MRiR) including all deaths in the follow-up period, but excluding deaths in the first 2 years of follow-up from all subsequent analyses, since CWP may be a manifestation of an existing illness. Starting with a basic model containing CWP, age group and sex, we added, individually, lifestyle factors or markers that could possibly mediate any observed relationship. We then added all such potential mediators to a final model. Associations are expressed MRiR with 95% CI.

Meta-analysis

We used (in a modified way) and updated a search conducted by Smith *et al*,[10] which identified studies examining the relationship between chronic pain and/or widespread pain (WP) and mortality. Although their review focused generally on chronic pain, our update focused only on studies examining WP or CWP. A second difference is that although previous meta-analyses extracted effect measures that were maximally adjusted for potential confounding factors, we have extracted data that are (as close as possible) only adjusted for age and sex. The difference is that we are answering the question 'Do persons with CWP experience excess mortality (in comparison to those without chronic pain)', whereas using fully adjusted effect measures is answering the question of whether the report of pain (per se) is associated with excess mortality. Thus the data on effect measures extracted from studies which they identified as eligible may be different.

We reran the published search strategy (in the appendix S1 of the original meta-analysis) from January 2014 (in order to ensure that articles published close to the time of the previous search were not missed) to January 2017.

Studies were eligible for the current meta-analysis provided that they:

- were observational studies
- sampled from a population sampling frame (or an approximation to such)
- identified persons with WP or CWP (including FM) and a comparison group of persons without such pain; the definition of WP should involve recognised criteria or the reporting of pain all over the body
- provided either a mortality rate ratio (MRtR) or MRiR quantifying the relationship between WP or CWP and mortality
- were published as a manuscript in English in a peer-reviewed journal.

Identified abstracts were screened by two authors and any disagreement resolved by discussion. We also checked studies included in the meta-analysis by Smith *et al*[10] to determine that they met the above eligibility criteria. Meta-analysis was conducted using a random effects model to reflect known differences in studies including geographical location, phenotypes and follow-up. The effect measures extracted from the eligible studies (MRrR or MRiR) were as closely as possible only adjusted for age and sex. In the meta-analysis, conducted using RevMan software, mortality risk ratio (MRR) has been used to signify the combined estimates using MRtR and MRiR. Sources of heterogeneity in effect measures were explored, specifically in relation to the geographical area in which the study was conducted and prevalence estimate of the phenotype studied.

RESULTS

UK Biobank

From 502 627 UK Biobank participants, 2193 (0.4%) did not answer the pain questions and are therefore excluded from this analysis. Among the remaining 500 434 persons, 7130 reported CWP (prevalence 1.4%), while 281 718 reported that they did not have any chronic pain. These two subcohorts are the study population for the current analysis, and their characteristics are shown in table 1. The CWP and the 'no chronic pain' groups had the same median age (58 years). Persons with CWP were less likely to be male (36.3% vs 50%); were more likely to be

Table 1 Characteristics of persons with chronic widespread pain (CWP) and no chronic pain in UK Biobank study

Characteristics	CWP (n=7130)	No chronic pain (n=281 718)
Died during follow-up (n, %)	405 (5.7%)	6493 (2.3%)
Died in first 2 years of follow-up (n,%)	95 (1.3%)	1224 (0.4%)
Age (median years, IQR)	58 (50, 63)	58 (52, 63)
Sex (% male)	2586 (36.3%)	135 186 (50.0%)
Body mass index (kg/m^2)		
Underweight (<18.5)	44 (0.6%)	1609 (0.6%)
Normal (18.5–24.9)	1354 (19.0%)	101 010 (35.9%)
Overweight (25.0–29.9)	2572 (36.1%)	121 141 (43.0%)
Obese (30.0–34.9)	1761 (24.7%)	43 088 (15.3%)
Obese (35.0–39.9)	799 (11.2%)	10 364 (3.7%)
Obese (≥40.0)	600 (8.4%)	4506 (1.6%)
Physical activity (mean min/week; SD)		
Walking	350 (579)	363 (511)
Moderate activity	276 (543)	270 (444)
Vigorous activity	72 (275)	93 (192)
Physical activity (climbing stairs per day)		
None	1223 (18.5%)	22 451 (8.1%)
1–5 times	2158 (32.6%)	53 163 (19.1%)
6–10 times	1795 (27.1%)	103 353 (37.2%)
11–15 times	720 (10.9%)	53 779 (19.4%)
16–20 times	378 (5.7%)	25 048 (9.0%)
>20 times	346 (5.2%)	20 071 (7.2%)
Smoking status (n,%)		
Current smoker	1316 (18.6%)	26 241 (9.3%)
Ex-regular smoker	1779 (25.1%)	61 161 (21.8%)
Ex-occasional smoker	627 (8.9%)	32 581 (11.6%)
Never or very rarely	3360 (47.4%)	160 839 (57.3%)
Diet: fruit and vegetable consumption (median portions/day, IQR)	8 (5, 11)	7 (5, 10)
Alcohol consumption (n, %)		
Daily or almost daily	767 (10.8%)	60 829 (21.6%)
3–4 times/week	842 (11.8%)	69 667 (24.7%)
1–2 times/week	1485 (20.9%)	74 096 (26.3%)
<1 time/week	2407 (33.8%)	58 139 (20.7%)
Never	1616 (22.7%)	18 789 (6.7%)

heavier than normal weight (80.4% vs 63.5%); be a current smoker (18.6% vs 9.3%); and not to drink any alcohol (22.7% vs 6.7%). They also undertook physical activity less often. In total there were 12 799 deaths in the study population within the period of observation: 7486 (58%) classified as being due to cancer, 2691 (21%) cardiovascular disease, 728 (6%) respiratory disease, 436 (3%) due to external causes and 1458 (11%) were classified as 'other'.

After adjusting for age and sex, participants with CWP had a more than twofold risk of dying in the follow-up period (MRiR 2.56, 95% CI 2.32 to 2.82), an excess that remained largely unchanged when deaths occurring in the first 2 years of follow-up were excluded (2.43, 95% CI 2.17 to 2.72). Deaths occurring in the first 2 years are excluded from all further analyses. Specific causes of death in excess were cancer ($1.73_{\text{adjusted age and sex}}$, 95% CI 1.46 to 2.05), cardiovascular ($3.24_{\text{adjusted age and sex}}$, 95% CI 2.55 to 4.11), respiratory ($5.66_{\text{adjusted age and sex}}$, 95% CI 4.00 to 8.03) and other disease-related causes ($4.04_{\text{adjusted age and sex}}$, 95% CI 3.05 to 5.34), while the excess of deaths from external causes was not statistically significant ($1.55_{\text{adjusted age and sex}}$, 95% CI 0.68 to 3.49).

We then examined to what extent the factors that were identified as being associated with pain status also predicted death in the period of follow-up (table 2). Age-adjusted risk of death was lower in women (MRiR 0.58 (95% CI 0.56 to 0.60). Age and gender adjusted risk was higher in obese participants (35–39 kg/m^2 vs normal weight 5.54 (95% CI 5.08 to 6.03), ≥40 kg/m^2 9.02 (95% CI 8.23 to 9.89)), and those who reported no walking (vs 1–100 min/week: 4.15 (95% CI 3.77 to 4.57)) or no moderate physical activity (vs 1–60 min/week: 2.95 (95% CI 2.74 to 3.19)). Risk of death was also higher in smokers (current smokers 2.54 (95% CI 2.39 to2.70), ex-smokers 1.44 (95% CI 1.36 to 1.52)) and persons who reported never drinking alcohol (vs daily drinkers 6.18 (95% CI 5.68 to 6.73)).

Finally, we tested to what extent adjusting the risk models for these measured lifestyle variables attenuated the relationship between CWP and excess mortality (table 3). Such attenuation would be consistent with the effects being mediated through such variable(s). When we did this, each class of variable (physical activity, BMI, smoking, diet including alcohol) when added to the model containing only pain status (CWP/no chronic pain), age group and sex resulted in a small attenuation of effect from an MRiR of 2.4 to MiRRs in the range of 2.0–2.2. However when all such potentially mediating variables were entered into the model, the MiRR reduced to 1.47 (95% CI 1.24 to 1.73). In cause-of-death-specific models with potential mediating variables, there remained an excess risk of cardiovascular 1.99 (95%

Table 2 Relationship between demographic and lifestyle factors and risk of death

Characteristics	Status at end of follow-up		Restricted model: mortality risk ratio (95% CI)*	Multivariable model: mortality risk ratio (95% CI)†
	Alive (n)	Dead (n)‡		
Pain status				
Chronic widespread pain	6725	310	2.43 (2.17 to 2.72)	1.47 (1.24 to 1.73)
No chronic pain	275 225	5269	Reference	Reference
Age group (years)				
<45	31 373	189	Reference	Reference
45–49	38 228	353	1.60 (1.37 to 1.87)	1.60 (1.25 to 2.07)
50–54	43 174	590	2.50 (2.17 to 2.89)	2.46 (1.95 to 3.11)
55–59	51 083	1021	3.80 (3.32 to 4.36)	3.61 (2.90 to 4.51)
60–64	67 884	2078	5.61 (4.92 to 6.39)	5.59 (4.51 to 6.92)
>64	50 538	2667	9.09 (7.98 to 10.4)	8.91 (7.20 to 11.0)
Sex				
Male	133 453	4319	Reference	Reference
Female	148 497	2579	0.58 (0.56 to 0.60)	0.59 (0.55 to 0.63)
Body mass index (kg/m^2)				
Underweight (<18.5)	1569	84	1.86 (1.40 to 2.50)	2.73 (2.07 to 3.60)
Normal (18.5–24.9)	100 295	2069	Reference	Reference
Overweight (25.0–29.9)	120 888	2825	1.70 (1.59 to 1.82)	0.93 (0.86 to 1.01)
Obese (30.0–34.9)	43 579	1270	3.20 (2.98 to 3.43)	1.11 (1.01 to 1.22)
Obese (35.0–39.9)	10 784	379	5.54 (5.08 to 6.03)	1.35 (1.16 to 1.58)
Obese (≥40.0)	4835	271	9.02 (8.23 to 9.89)	1.94 (1.59 to 2.36)
Physical activity: walking (min/week)				
0	5150	225	4.15 (3.77 to 4.57)	1.19 (0.99 to 1.43)
1–100	63 711	1547	Reference	Reference
101–210	74 315	1778	0.73 (0.68 to 0.79)	0.98 (0.90 to 1.07)
211–420	58 945	1312	0.64 (0.59 to 0.69)	0.92 (0.83 to 1.01)
>420	46 710	1017	0.85 (0.79 to 0.92)	0.89 (0.79 to 0.99)
Physical activity: moderate (min/week)				
0	32 562	1127	2.95 (2.74 to 3.19)	1.14 (1.02 to 1.27)
1–60	60 247	1221	Reference	Reference
61–150	51 037	1086	0.91 (0.83 to 0.99)	0.99 (0.90 to 1.10)
151–360	51 640	1086	0.87 (0.79 to 0.95)	0.98 (0.88 to 1.10)
>360	49 171	1229	1.30 (1.20 to 1.42)	1.09 (0.97 to 1.22)
Physical activity: vigorous (min/week)				
0	94 509	3068	Reference	Reference
1–40	45 581	915	0.37 (0.34 to 0.40)	0.77 (0.69 to 0.85)
41–90	40 814	729	0.30 (0.28 to 0.33)	0.78 (0.70 to 0.87)
91–180	39 355	678	0.27 (0.24 to 0.30)	0.76 (0.68 to 0.85)
>180	33 648	645	0.43 (0.39 to 0.47)	0.79 (0.70 to 0.89)
Physical activity: stairs (times/day)				
0	22 789	885	1.29 (1.20 to 1.38)	1.02 (0.91 to 1.14)
1–5	53 707	1614	Reference	Reference
6–10	102 928	2220	0.43 (0.41 to 0.46)	0.83 (0.76 to 0.91)
11–15	53 420	1079	0.33 (0.30 to 0.36)	0.86 (0.77 to 0.95)
16–20	24 986	440	0.37 (0.33 to 0.41)	0.69 (0.60 to 0.80)
>20	20 011	406	0.42 (0.38 to 0.47)	0.92 (0.80 to 1.07)
Smoking status				
Current smoker	26 309	1248	2.54 (2.39 to2.70)	2.31 (2.10 to 2.54)
Ex-regular smoker	60 770	2170	1.44 (1.36 to 1.52)	1.55 (1.43 to 1.67)
Ex-occasional smoker	32 532	676	0.92 (0.85 to 1.003)	1.09 (0.97 to 1.22)
Never or very rarely	161 432	2767	Reference	Reference
Alcohol consumption				
Daily or almost daily	59 954	1642	Reference	Reference

Continued

Table 2 Continued

Characteristics	Status at end of follow-up: Alive (n)	Status at end of follow-up: Dead (n)‡	Restricted model: mortality risk ratio (95% CI)*	Multivariable model: mortality risk ratio (95% CI)†
3–4 times/week	69 132	1377	0.96 (0.87 to 1.06)	0.92 (0.84 to 1.02)
1–2 times/week	73 949	1632	1.57 (1.44 to 1.72)	0.97 (0.88 to 1.07)
<1 time/week	59 073	1473	3.08 (2.84 to 3.34)	1.08 (0.98 to 1.20)
Never	19 639	766	6.18 (5.68 to 6.73)	1.49 (1.32 to 1.69))
Diet: fruit and vegetable consumption				
Lowest consumption	62 641	1802	Reference	Reference
Quintile 2	58 079	1363	0.74 (0.69 to 0.80)	0.90 (0.82 to 0.98)
Quintile 3	25 448	569	0.78 (0.71 to 0.86)	0.88 (0.78 to 0.99)
Quintile 4	50 750	1156	0.75 (0.69 to 0.80)	0.88 (0.80 to 0.97)
Highest consumption	40 733	881	1.01 (0.94 to 1.09)	0.86 (0.77 to 0.95)

*Adjusted for age and/or sex as applicable and excluding first 2 years of follow-up.
†All variables entered into the statistical model and mutually adjusted.
‡Deaths within 2 years of the baseline assessment are excluded.

CI 1.41 to 2.80), respiratory 1.91 (95% CI 1.08 to 3.36) and 'other disease' deaths 2.14 (95% CI 1.42 to 3.21), but the there was no longer an excess risk of cancer death 1.06 (95% CI 0.82 to 1.38) and external deaths 1.01 (95% CI 0.30 to 3.40).

Meta-analysis

Our search identified 3171 unique publications, of which 15 proceeded to abstract screening and 1 to full-text screening and subsequent inclusion.[12] Of the five studies included in the meta-analysis of Smith *et al*,[10] one did not meet eligibility criteria for the current meta-analysis,[13] since the pain phenotype did not include any measure of 'widespreadness'. Instead the phenotype examined was multiple joint pain. Thus a total of six studies (including the current analysis) were eligible for the current meta-analysis.[1 3 12 14 15] Characteristics of studies identified as eligible are given in table 4. One study presented data only to one decimal place and thus in the meta-analysis was identified as having a non-symmetrical log-transformed CI.[3] We therefore contacted the first author of the publication and they provided more precise data (for analyses only adjusted for age and sex). Eligible studies included 580 020 participants from three European countries (Norway, Sweden and the UK). There was significant heterogeneity between studies: I^2=98% for all-cause mortality, 95% for cardiovascular, 96% for respiratory and 91% for cancer (all p<0.001). All studies showed significant excess of all-cause mortality and the combined estimate of this was 57% (MRR 1.57; 95% CI 1.06 to 2.33). For cardiovascular mortality, three out of five studies showed a significant association and the combined estimate of this was 63% (1.63; 95% CI 0.98 to 2.70). For respiratory mortality, only one out of three studies showed a significant excess mortality, and there was considerable uncertainty around the pooled estimate of excess risk (1.70; 95% CI 0.45 to 6.45). For cancer, three out of five studies showed significant excess mortality and the pooled estimate was 51% (1.51; 95% CI 1.06 to 2.13) (figure 1).

We investigated the source of heterogeneity with respect to the relationship between CWP and all-cause mortality. When restricted by geographical area, the meta-analysis showed that considerable heterogeneity was present in studies conducted in Great Britain (I^2=90%) (MRR 1.60; 95% CI 1.06 to 2.42) but not

Table 3 Relationship between pain status and risk of death, adjusting for potential mediating variables

Variables added to basic model*	Participants† included in model (N)	MRR‡ (95% CI) CWP versus no chronic pain	MRR (95% CI) CWP versus no chronic pain (participants with full data)§
No additional variables	287 529	2.43 (2.17 to 2.72)	2.23 (1.90 to 2.62)
Body mass index category¶	287 529	2.13 (1.90 to 2.39)	1.98 (1.68 to 2.33)
Physical activity: walking	253 579	2.09 (1.82 to 2.40)	2.08 (1.76 to 2.44)
Physical activity: moderate	249 309	2.23 (1.96 to 2.54)	2.06 (1.75 to 2.42)
Physical activity: vigorous	258 755	2.22 (1.97 to 2.51)	2.01 (1.71 to 2.36)
Physical activity: stairs	283 221	2.12 (1.88 to 2.38)	2.07 (1.76 to 2.43)
Smoking	286 590	2.16 (1.94 to 2.42)	2.01 (1.71 to 2.37)
Diet: alcohol consumption	287 320	2.21 (1.97 to 2.47)	2.05 (1.74 to 2.41)
Diet: fruit and vegetables	242 346	2.30 (2.02 to 2.60)	2.21 (1.88 to 2.60)
Full multivariable model**	193 676	1.47 (1.24 to 1.73)	1.47 (1.24 to 1.73)

**All additional variables entered into model: age, sex, body mass index, physical activity (walking, moderate and vigorous activities, climbing stairs), diet (fruit and vegetable, alcohol consumption) and smoking status.
†Deaths occurring within 2 years of the baseline assessment are excluded.
‡Mortality risk ratio.
§Restricted to 193 676 participants with data on all variables included in the full model.
¶Each line represents the basic model with the addition of the single variable stated.
CWP, chronic widespread pain.

Table 4 Studies eligible for meta-analysis of CWP and mortality

Study (location)	Sampling frame	Pain phenotype	Pain phenotype prevalence (%)	Deaths (n)/Study (n)	Follow-up (years)
Andersson[15]	Random sample in two municipalities	>4 pain locations representing both the upper and lower body and including axial pain	9.4	189/1609	14
Åsberg *et al*[11]	All inhabitants of one county	CWP modified* definition in ACR 1990 criteria of FM	23.1	12 521/65 026	14
Macfarlane *et al*[1]	Persons registered with GP in two areas	WP according to definition in ACR 1990 criteria of FM	15.3	654/6569	8
Macfarlane *et al* (current study)	Persons aged 40–69 registered with GP in 22 areas	'Pain all over the body' lasting ≥3 months	1.4	12 799/288 848	7
McBeth *et al*[3]	Age-stratified and sex-stratified sample from 3 GPs in one region	WP definition in ACR 1990 criteria of FM	16.9	1017/4344	8
Nitter and Forseth[14]	Women born in 1940–1969 in one town	Pain in muscles and joints and back, or pain in whole body, lasting ≥3 months	12.9	89/2038	18

*There was no requirement to have pain on both sides of the body.
ACR, American College of Rheumatology; CWP, chronic widespread pain; GP, general practitioner; FM,fibromyalgia; WP, widespread pain.

in studies conducted in Scandinavia (I^2=0%) (MRR 1.06; 95% CI 1.02 to 1.10). Similarly when analysis was restricted to those studies with prevalence of CWP in the 10%–20% mid-range, that is, excluding those with the extreme prevalence estimates, there was no evidence of heterogeneity (I^2=0%) (MRR 1.30; 95% CI 1.07 to 1.58).

DISCUSSION

Using data from UK Biobank, involving over half a million study participants, we have demonstrated that persons with CWP have an important excess of risk of dying in the medium and long term. This excess risk was evident across all disease and non-disease categories. The meta-analysis of this relationship shows that all six studies conducted find excess mortality and estimate the excess risk across all studies at 59%, although there is significant heterogeneity. Similar excesses of cancer and cardiovascular mortality are observed. In UK Biobank, adjustment for lifestyle factors substantially reduced the excess risk, and this observation is consistent with them mediating the relationship between CWP and mortality.

Methodological issues

The main strengths of UK Biobank in addressing this question include that it uses a sampling frame that is considered to have almost complete population coverage. Although the participation rate was low (5.5%), we have previously published an analysis that demonstrates that the prevalence of regional pains in UK Biobank is very similar to more traditional pain epidemiological studies with higher participation, and that the study reproduces known relationships with aetiological factors.[16] The large sample has allowed us to examine specific causes of death to exclude deaths within 2 years of the assessment (since WP may be a manifestation of a disease linked to death, eg, metastatic cancer) and consider the role of mediating factors.

The phenotype used in studies that have examined the relationship with mortality has varied considerably. They have included WP according to the definition within the American College of Rheumatology (ACR) criteria (1990) for FM,[1 3] and modifications of the ACR 1990 FM criteria in terms of pain timing and distribution[12 15] or bespoke definitions to capture 'widespreadness'.[14] The comparison populations also differ: persons who are free of pain,[1 3 14] free of chronic pain[15] or who simply do not meet the phenotype[12] are variously used. Some studies had an additional criterion that WP is required to be chronic, although studies of WP have shown that the vast majority of persons with WP report chronic symptoms (81% in UK Biobank). These have resulted in prevalence proportions within population-based studies of between 1.4% and 23.1% and suggest important differences in the symptomatic populations studied. Interestingly the study with the highest prevalence[12] reported a markedly lower excess risk of mortality. UK Biobank has used the most stringent definition, which has resulted in a prevalence similar to that of FM,[17] and across all-cause and disease-specific mortality reports some of the highest excess mortality. This is consistent with the hypothesis that the greatest excess mortality is among those with more severe symptoms. Sensitivity analyses confirmed that heterogeneity in risk estimates was indeed partly explained by differences in prevalence, as well as by geographical area.

We have approached the analysis in a different way to some previous studies on this topic. We adjusted for the confounding factors of age and sex. Given that the question we are asking is 'Do patients with CWP experience prematurely mortality?' we believed that no further adjustment should be made. However when excess mortality is observed, it is of relevance to examine mediators — since these can become targets for intervention. Previous studies have identified lack of physical activity and poor quality diet as the variables that may explain a relationship. UK Biobank has a rich source of data to allow the assessment of these potential mediators. They nevertheless represent markers of these lifestyle factors rather than comprehensive assessments. Despite this, adjustment for these lifestyle markers almost completely explained cancer and 'non-disease' excess mortality and explained 56%, 80% and 62% of the excess mortality for cardiovascular, respiratory and 'other-disease', respectively.

Comparison with other studies and coherence of evidence

UK Biobank has provided results that are generally consistent with previously conducted studies. For cardiovascular mortality it has provided the largest estimate for excess mortality. It is the first study to suggest a relationship between CWP and excess mortality from respiratory disease.

The meta-analysis of Smith *et al*[10] conducted on this topic chose to extract the most fully adjusted model available in included studies, which means that this examines a subtly different question of whether pain directly increases mortality risk (independent of any lifestyle, psychosocial or clinical factors). We believe

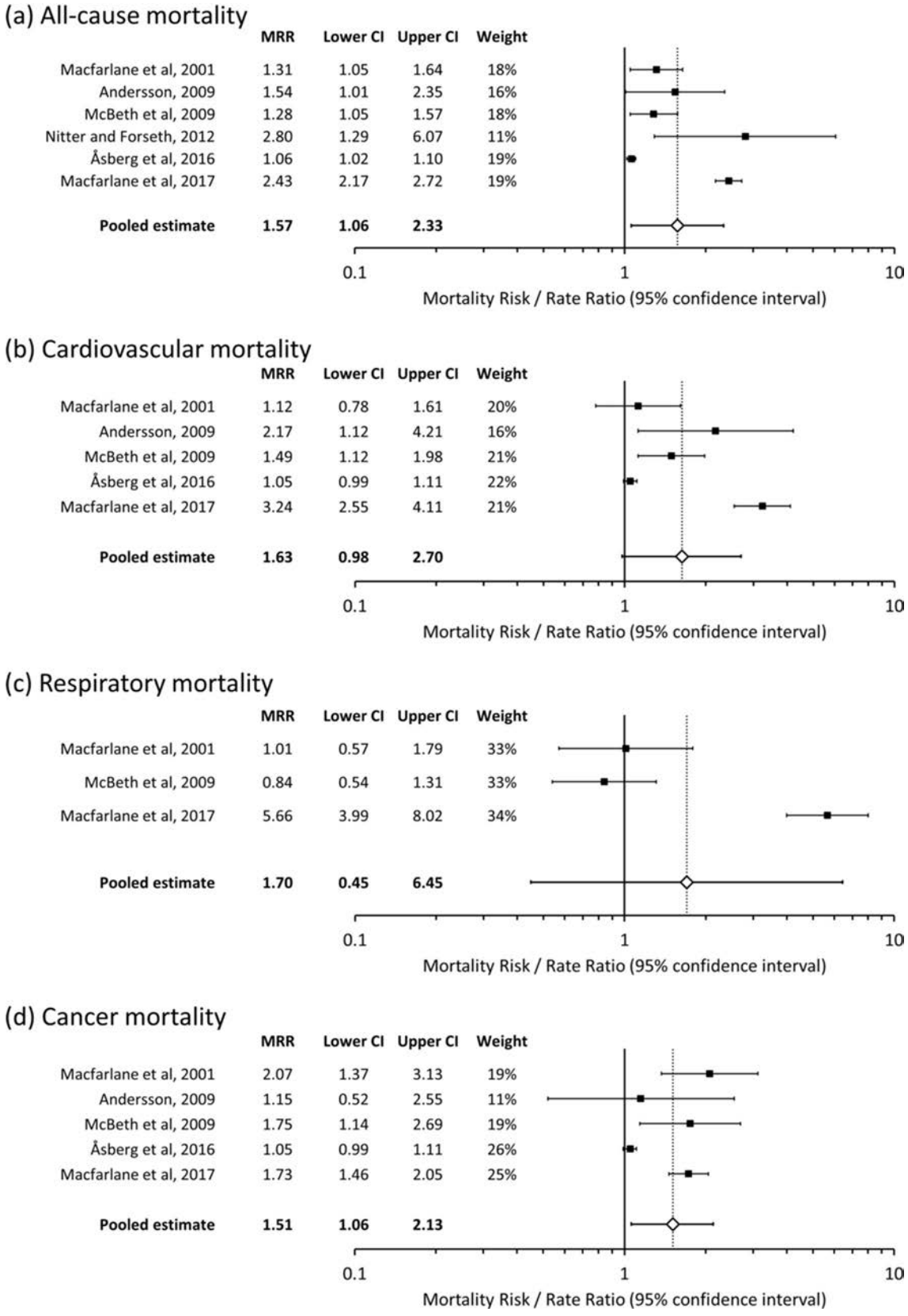


Figure 1 Forest plots of pain and all-cause and disease-specific mortality. MRR, mortality risk ratio.

that the most clinically relevant question for clinicians managing patients with WP/CWP or FM is what factors can be modified that could reduce any excess mortality which such patients experience. We also excluded one study included in the previous meta-analysis. The study of Macfarlane *et al*[13] was not eligible for this analysis as it examined the mortality consequences of multijoint pain (at least four joints). There was no requirement for pain to be widespread. All included studies had some

requirement for the pain to be widespread or for the participant to endorse that the pain was all over their body. Even if the study of Macfarlane *et al*,[13] which did not find any excess mortality MRiR (0.86; 0.74,1.01), had been included in the meta-analysis, the combined estimate would still have suggested an important excess. Exclusion of a phenotype that excludes a measure of 'widespreadness' is supported by a proposed modification to the 2011 research criteria for FM, which requires that multisite pain is also widespread across the body.[18] The meta-analysis of Åsberg *et al*[19] concluded that 'pooled data gave no evidence for a higher mortality rate among individuals with chronic widespread musculoskeletal complaints'. This put emphasis on a pooled unadjusted MRR of 1.69, which was not statistically significant, and a markedly reduced excess (MRR 1.13) after full adjustment. The inclusion of UK Biobank, considering age-adjusted and sex-adjusted risks, has provided a similar pooled estimate of excess risk (MRR 1.59) and is now statistically significant.

We conclude that the evidence is now clear that persons with CWP experience excess mortality. UK Biobank results considerably reduce uncertainty around the magnitude of excess risk, and demonstrate that the risk is unlikely to be due to the experience of pain per se, but is substantially explained by lifestyle factors associated with having pain (poor diet, low levels of physical activity, smoking, high BMI). These provide important targets for intervention in managing patients with CWP. Optimal management of FM should include exercise, but this is often not provided in a structured and supported way to facilitate long-term behaviour change. Few patients with CWP or FM receive specific supported care in improving diet or stopping smoking. The data from this study show that changing the habits of persons with CWP to be similar to persons without CWP could reduce mortality by around 35%. Such approaches should have high priority in the routine care of such patients.

Acknowledgements This manuscript uses the UK Biobank resource (Application 1144). We acknowledge the authors of a previous meta-analysis on this topic (Diane Smith, Ross Wilkie, Olalekan Uthman, Joanne L Jordan, John McBeth) whose published search strategy we used as the basis for our meta-analysis, although our meta-analysis had a more restricted focus and the criteria for determining eligibility and the data we extracted from eligible studies were not identical and resulted in selection of a different group of studies. We thank John McBeth (University of Manchester) for providing additional data relating to one of the studies, to allow it be included in the meta-analysis.

Contributors GJM had the idea for the study and together with GTJ designed the analysis plan for UK Biobank. GTJ undertook the UK Biobank analysis. MSB conducted the updated systematic review, and all authors participated in undertaking the meta-analysis. GJM drafted the manuscript, but all authors made an important intellectual contribution to the text.

Competing interests None declared.

Ethics approval North West Multi-centre Research Ethics Committee (MREC).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Macfarlane GJ, McBeth J, Silman AJ. Widespread body pain and mortality: prospective population based study. *BMJ* 2001;323:662–5.
2 McBeth J, Silman AJ, Macfarlane GJ. Association of widespread body pain with an increased risk of cancer and reduced cancer survival: a prospective, population-based study. *Arthritis Rheum* 2003;48:1686–92.
3 McBeth J, Symmons DP, Silman AJ, *et al*. Musculoskeletal pain is associated with a long-term increased risk of Cancer and cardiovascular-related mortality. *Rheumatology* 2009;48:74–7.
4 Kyu HH, Bachman VF, Alexander LT, *et al*. Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013. *BMJ* 2016;354:i3857.
5 Liu Y, Hu F, Li D, *et al*. Does physical activity reduce the risk of prostate cancer? A systematic review and meta-analysis. *Eur Urol* 2011;60:1029–44.
6 McBeth J, Nicholl BI, Cordingley L, *et al*. Chronic widespread pain predicts physical inactivity: results from the prospective EPIFUND study. *Eur J Pain* 2010;14:972–9.
7 Mundal I, Gråwe RW, Bjørngaard JH, *et al*. Prevalence and long-term predictors of persistent chronic widespread pain in the general population in an 11-year prospective study: the HUNT study. *BMC Musculoskelet Disord* 2014;15:213.
8 Mundal I, Gråwe RW, Bjørngaard JH, *et al*. Psychosocial factors and risk of chronic widespread pain: an 11-year follow-up study--the HUNT study. *Pain* 2014;155:1555–61.
9 Vandenkerkhof EG, Macdonald HM, Jones GT, *et al*. Diet, lifestyle and chronic widespread pain: results from the 1958 british Birth Cohort Study. *Pain Res Manag* 2011;16:87–92.
10 Smith D, Wilkie R, Uthman O, *et al*. Chronic pain and mortality: a systematic review. *PLoS One* 2014;9:e99048.
11 Åsberg AN, Stovner LJ, Zwart JA, *et al*. Chronic musculoskeletal complaints as a predictor of mortality-The HUNT study. *Pain* 2016;157:1443–7.
12 Sudlow C, Gallacher J, Allen N, *et al*. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
13 Macfarlane GJ, Jones GT, Knekt P, *et al*. Is the report of widespread body pain associated with long-term increased mortality? data from the Mini-Finland Health Survey. *Rheumatology* 2007;46:805–7.
14 Nitter AK, Forseth Karin. Mortality rate and causes of death in women with self-reported musculoskeletal pain: Results from a 17-year follow-up study. *Scandinavian Journal of Pain* 2013;4:86–92.
15 Andersson HI. Increased mortality among individuals with chronic widespread pain relates to lifestyle factors: a prospective population-based study. *Disabil Rehabil* 2009;31:1980–7.
16 Macfarlane GJ, Beasley M, Smith BH, *et al*. Can large surveys conducted on highly selected populations provide valid information on the epidemiology of common health conditions? An analysis of UK Biobank data on musculoskeletal pain. *Br J Pain* 2015;9:203–12.
17 Jones GT, Atzeni F, Beasley M, *et al*. The prevalence of Fibromyalgia in the general population: a comparison of the American College of Rheumatology 1990, 2010, and modified 2010 classification criteria. *Arthritis Rheumatol* 2015;67:568–75.
18 Åsberg AN, Heuch I, Hagen K. The Mortality Associated With Chronic Widespread Musculoskeletal Complaints: A Systematic Review of the Literature. *Musculoskeletal Care* 2017;15:104–13.
19 Wolfe F, Clauw DJ, Fitzcharles MA, *et al*. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. *Semin Arthritis Rheum* 2016;46:319–29.

EXTENDED REPORT

Sacroiliac radiographic progression in recent onset axial spondyloarthritis: the 5-year data of the DESIR cohort

Maxime Dougados,[1,2] Alexandre Sepriano,[3,4] Anna Molto,[2,5] Miranda van Lunteren,[3] Sofia Ramiro,[3] Manouk de Hooge,[3] Rosaline van den Berg,[3] Victoria Navarro Compan,[6] Christophe Demattei,[7] Robert Landewé,[8] Désirée van der Heijde[3]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211596).

[1]Department of Rheumatology, Paris Descartes University, Hôpital Cochin, Hôpitaux de Paris, Paris, France
[2]INSERM (U1153), Clinical Epidemiology and Biostatistics, PRES Sorbonne Paris-City, Paris, France
[3]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[4]NOVA Medical School, Universidade Nova de Lisboa, Lisbon, Portugal
[5]Department of Rheumatology, Hôpital Cochin, Paris, France
[6]Department of Rheumatology, University Hospital La Paz, IdiPaz, Madrid, Spain
[7]Department of Biostatistics, Epidemiology, Public Health and Medical Information (BESPIM), Nimes University Hospital, Nimes, France
[8]Amsterdam Rheumatology & Clinical Immunology Center (ARC), Amsterdam and Zuyderland Medical Center, Heerlen, The Netherlands

Correspondence to
Professor Maxime Dougados, INSERM (U1153), Clinical Epidemiology and Biostatistics, PRES Sorbonne Paris-City, Paris 14, France; maxime.dougados@aphp.fr

Received 4 April 2017
Revised 5 June 2017
Accepted 5 June 2017
Published Online First 20 July 2017

To cite: Dougados M, Sepriano A, Molto A, *et al*. *Ann Rheum Dis* 2017;**76**:1823–1828.

ABSTRACT

Objective To estimate sacroiliac joint radiographic (X-SIJ) progression in patients with axial spondyloarthritis (axSpA) and to evaluate the effects of inflammation on MRI (MRI-SIJ) on X-SIJ progression.

Methods X-SIJ and MRI-SIJ at baseline and after 2 and 5 years in patients with recent onset axSpA from the DESIR cohort were scored by three central readers. Progression was defined as (1) the shift from non-radiographic (nr) to radiographic (r) sacroiliitis (by modified New York (mNY) criteria) or alternative criteria, (2) a change of at least one grade or (3) a change of at least one grade but ignoring a change from grade 0 to 1. The effects of baseline inflammation on MRI-SIJ on 5-year X-SIJ damage (mNY) were tested by generalised estimating equations.

Results In 416 patients with pairs of baseline and 5-year X-SIJ present, net progression occurred in 5.1% (1), 13.0% (2) and 10.3% (3) respectively, regarding a shift from nr-axSpA to r-axSpA (1), a change of at least one grade (2) or a change of at least one grade but ignoring a change from grade 0 to 1 (3). Baseline MRI-SIJ predicted structural damage after 5 years in human leukocyte antigen-B27 (HLA-B27) positive (OR 5.39 (95% CI 3.25 to 8.94)) and in HLA-B27 negative (OR 2.16 (95% CI 1.04 to 4.51)) patients.

Conclusions Five-year progression of X-SIJ damage in patients with recent onset axSpA is limited but present beyond measurement error. Baseline MRI-SIJ inflammation drives 5-year radiographic changes.

INTRODUCTION

Axial spondyloarthritis (axSpA) comprises two subcategories based on the presence of structural changes in the sacroiliac joints (SIJs): radiographic (r)-axSpA and non-radiographic (nr)-axSpA. R-axSpA implies the fulfilment of the modified New York criteria (mNY).[1 2]

Information about the natural course of radiographic sacroiliitis and factors that contribute to it is scarce.[3] Prospective cohorts should give resolution, and long-term follow-up of patients with recent onset disease is mandatory to 'capture' meaningful progression. Inherently, such studies face the risk of loss to follow-up and attrition bias.

DESIR (acronym in French for outcome of recent onset spondyloarthritis) is a prospective cohort of patients with recent onset axSpA (NCT01648907). With this study, we address the primary objectives of DESIR, formulated as follows: (1) what proportion of patients switches from nr-axSpA to r-axSpA after 5 years?; (2) how sensitive are different outcome measures for radiographic damage of SIJ (X-SIJ) to change?; (3) does inflammation on MRI of the SIJ (MRI-SIJ) lead to structural damage on X-SIJ after 5 years?

METHODS

Patients

The DESIR cohort has been previously described.[4] Briefly, consecutive patients (aged 18–50 from 25 centres in France) with inflammatory back pain[5 6] and a duration ≥3 months but <3 years were included if the treating rheumatologist considered the symptoms suggestive of axSpA (a score ≥5 on a scale from 0 to 10, in which 0 was 'not suggestive' and 10 'very suggestive'). Between December 2007 and April 2010, 708 patients were included.

The study was conducted according to good clinical practice guidelines and was approved by the appropriate local medical ethical committees. A detailed description of the study protocol is available at the DESIR website (http://www.lacohortedesir.fr/desir-in-english/). The research proposal for this particular analysis was approved by the scientific committee of the DESIR cohort.

Clinical data

By using a standardised case report form (CRF) information was collected with questionnaires, physical examination, ongoing treatments and laboratory tests according to the DESIR protocol. The database used for this analysis was locked in June 2016.

At baseline, age, gender, smoking status, HLA-B27 and duration of axial symptoms had been collected. At baseline, every 6 months during the first 2 years of follow-up, and annually thereafter the following parameters had been collected: Bath Ankylosing Spondylitis Disease Activity Index (BASDAI),[7] Bath Ankylosing Spondylitis Functional Index,[8] C-reactive protein (CRP), treatment including non-steroidal anti-inflammatory drugs (NSAID) by the Assessment of Spondyloarthritis International Society (ASAS)-NSAID score and tumour necrosis factor inhibitors (TNFi).[9]

Pelvic radiographs

Pelvic radiographs collected at baseline, 2 years and 5 years of follow-up were evaluated in one session independently by three central readers (MdH,

VNC and RvdB). Readers were blinded for time order and clinical information. Each reader evaluated each SIJ according to the mNY grading method (0: normal; 1: suspicious changes; 2: minimal abnormalities; 3: unequivocal abnormalities; and 4: severe abnormalities (complete ankylosis).[10]

Pelvic MRI

MRI-SIJ collected at baseline, 2 years and 5 years of follow-up were evaluated in one session independently by three central readers (MdH, VNC and MvL). Readers were blinded for time order and clinical information. MRI-SIJ was considered positive if bone marrow oedema (BMO) lesions highly suggestive of SpA were present (either one BMO lesion on ≥2 consecutive slides or several BMO lesions on one slice).[11] An MRI-SIJ was considered positive if at least two out of three readers judged positivity. MRI-SIJ and X-SIJ were scored entirely independently.

Sample size calculation

The sample size calculation was based on an estimated prevalence of radiographic damage between 70% and 90% at year 5 irrespective of the baseline status. Moreover, we estimated the prevalence of inflammation on MRI-SIJ at baseline between 30% and 50%.[12 13]

The number of patients was calculated based on a relative risk of 2–3 to observe radiographic damage at year 5 in case of a baseline MRI-SIJ inflammation. For a 5% bilateral alpha risk, a 90% power, and the different assumptions including an attrition rate between 15% and 20%, the number of required patients ranged from 685 to 768, and 700 was the chosen number.

Statistical analysis

SIJ radiographic progression

The 5-year X-SIJ progression was assessed in patients in whom baseline and year 5 X-SIJ were present (completers' population). Assessed were: (A) switch from nr-axSpA at baseline to r-axSpA (mNY score) at 5 years; (B) worsening of at least one grade in at least one SIJ; (C) worsening of at least one grade in at least one SIJ, but with a 5-year grade of at least 2 in the worsened joint; and (D) change in the total mNY score (expressed as a continuous variable) with a range from 0 to 8 (4 grades per SIJ).

In order to give sufficient credit to measurement error, we determined the proportion of 'progressors' (% of patients with worsening) as well as the proportion of 'regressors' (% of patients with improvement). Improvement was defined per outcome measure: (A) switching from r-axSpA at baseline to nr-axSpA at 5 years; (B) reduction of at least one grade in at least one SIJ; and (C) reduction of at least one grade in at least one SIJ with a baseline score of at least 2 in the improved joint. In addition, 'net' percentage of progression was defined as the number of 'progressors' minus the number of 'regressors' (numerator) divided by the total number of the study population (denominator) and was analysed in the entire population and clinically relevant subgroups.

Sensitivity analyses that addressed the impact of missing data were performed in patients with a baseline and at least one postbaseline radiograph available ('intention-to-follow' population) using two imputation techniques:(1) last observation carried forward (LOCF) and (2) linear extrapolation (LE).

The continuous SIJ score (total scores of left plus right SIJ (ranging from 0 to 8)) was the mean score of the three readers; for the binary definitions, a change was considered present if at least two out of the three readers agreed.

Effect of baseline MRI-SIJ inflammation on the 5-year X-SIJ damage

The association between baseline MRI-SIJ inflammation and 5-year X-SIJ damage (primary outcome) was analysed by three different models: (1) binomial multivariable generalised estimating equations (GEEs) on the individual readers' scores (1-level GEE model); (2) 'traditional' multivariable logistic regression on the aggregated (two out of three reader consensus scores for MRI and SIJ) X-SIJ progression scores; (3) a true longitudinal (2-level) multivariable GEE with time-lagged autoregressive variables (as in Ramiro *et al*).[14] The logistic regression models were also fit after multiple imputations with chainedequations (MICE) in the 'intention-to-follow' population.

Potential baseline confounders for the association of interest were selected based on their clinical relevance (gender, symptom duration, CRP, BASDAI, smoking status and treatment with NSAIDs). Statistical interactions between MRI-SIJ inflammation and baseline variables were excluded first and, if relevant ($p<0.15$ for the interaction term), the model was fitted per stratum.

RESULTS

Patients and study course

Pelvic radiographs were available for 685 of the 708 patients at baseline. Of the 685 patients with baseline X-SIJ, 519 and 416 patients had X-SIJ, from all readers, after 2 and 5 years, respectively (completer's population). A postbaseline X-SIJ (either at year 2 or 5) was available for 557 patients (intention to follow population). A baseline MRI-SIJ was available for 679 patients.

Table 1 summarises the baseline characteristics for patients with complete 5-year pelvic radiograph data and those without.

Radiographic progression after 5 years of follow-up

At baseline, the mNY criteria were fulfilled by 62/416 (14.9%; according to two out of three readers) of the patients in the completers' population. After 5 years, this proportion has increased to 20.0% in the completers' population and to 18.0% and 17.7% in the 'intention-to-follow' population (n=557), after LOCF and LE, respectively. A statistically significant worsening of the mean (SD) SIJ score was found in all scenarios (from 1.41 (1.68) to 1.60 (1.83) (Δ:0.19 (0.55); $p<0.001$) in the completers' population and from 1.32 (1.65) to 1.49 (1.81) (Δ:0. 17 (0.59); $p<0.001$) (LOCF) or from 1.33 (1.65) to 1.50 (1.84) (Δ:0.17 (0.61); $p<0.001$) (LE) in the 'intention to follow' population).

Figure 1 summarises the observed changes in the binary outcome measures in the completers' population, in terms of '% worsened', '% improved' and 'net % progression' (online supplementary figures S1 and S2 provide the same information for the 'intention-to-follow' population after LOCF and LE, yielding similar results).

Effects of MRI-SIJ inflammation on X-SIJ damage

Figure 2 shows the effect of baseline MRI-SIJ inflammation on 5-year SIJ damage according to the mNY criteria, stratified for HLA-B27 (interaction: $p=0.033$). Baseline MRI-SIJ inflammation was associated with radiographic damage after 5 years in HLA-B27 positive patients (OR 5.39 (95% CI 3.25 to 8.94)) as well as HLA-B27 negative patients (OR 2.16 (95% CI 1.04 to 4.51)). The association between baseline MRI inflammation and 5-year SIJ damage was consistently found, regardless of the analytical method and the definition of SIJ progression (table 2).

Table 1 Baseline characteristics according to the availability of complete 5-year radiographic data of the sacroiliac joints

Characteristics	Status at year 5		
	Completers*	Non-completers	All patients
Number of patients	417	291	708
Age (mean, SD)	34.1 (8.6)	33.2 (8.6)	33.7 (8.6)
Symptom duration (years), (mean, SD)	1.5 (0.9) (n=416)	1.5 (0.8) (n=291)	1.5 (0.9) (n=707)
Male gender (%)	198 (47.5)	129 (44.3)	327 (46.2)
HLA-B27 positivity (%)	267 (64.0) (n=417)	143 (49.3) (n=290)	410 (58.0) (n=707)
X-SIJ structural damage† (mNY) (%)	62 (14.9) (n=416)	29 (10.8) (n=268)	92 (13.5) (n=684)
MRI-SIJ inflammation†‡ (%)	113 (28.1) (n=402)	67 (24.2) (n=277)	180 (26.5) (n=679)
Abnormal CRP§ (%)	126 (31.5) (n=400)	78 (27.4) (n=285)	204 (29.8) (n=685)
BASDAI (0–10, mean, SD)	4.34 (1.99) (n=416)	4.65 (2.01) (n=288)	4.47 (2.00) (n=704)
ASDAS (mean, SD)	2.6 (1.0) (n=395)	2.6 (0.9) (n=281)	2.6 (1.0) (n=676)
BASFI (0–10, mean, SD)	2.92 (2.24) (n=413)	3.23 (2.32) (n=288)	3.04 (2.28) (n=701)

*Patients with both baseline and 5-year X-SIJ available.
†According to the '2 out of 3' definition: agreement of at least two out of the three readers—if two readers disagree and the third reading is missing, the combined score is set as missing (one case for X-SIJ).
‡Presence of bone marrow oedema according to the ASAS criteria at MRI-SIJ.
§≥6 mg/L.
ASDAS, Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; mNY, modified New York criteria; MRI-SIJ, MRI of the sacroiliac joints; X-SIJ, radiograph of the sacroiliac joints.

Radiographic progression across clinically relevant subgroups

Figure 3 shows the 'net' progression from nr-axSpA to r-axSpA in different subgroups of patients according to relevant clinical characteristics and the interaction with HLA-B27. HLA-B27-positive nr-axSpA patients with a positive MRI-SIJ and CRP had a likelihood of 'net' progression of at least 1-grade of the X-SIJ mNY score that was more than twice as high as r-axSpA patients with similar baseline features (see online supplementary figures S3 and S4).

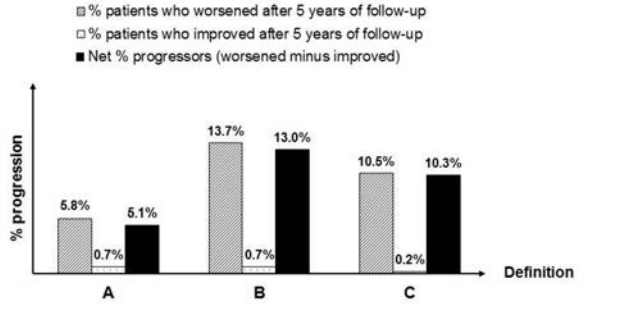
DISCUSSION

The main findings of this 5-year follow-up study can be summarised as follows: (1) 5-year radiographic SIJ progression is statistically significant but of limited magnitude; (2) strategically chosen definitions of radiographic progression may be more sensitive to change over time than the rigid (binary) mNY-based definition; and (3) inflammation on MRI-SIJ is highly predictive of a structural radiographic SIJ progression. Moreover, these data provide meaningful information for the clinician who likes to determine the risk of progression in an individual patient, using baseline parameters such as HLA-B27 positivity, radiographic structural damage, MRI-SIJ inflammation and abnormal CRP.

In order to properly interpret the rate of progression of SIJ damage that we found in this study, two quantities have to be considered: (A) the proportion of patients with radiographic SIJ damage at baseline; and (B) the proportion of patients that change from nr-axSpA to r-axSpA over time.

Observed radiographic SIJ damage in the DESIR cohort (15%) is in accordance with what has been found before, in light of the relatively short duration of the symptoms (between 3 months and 3 years).[15–17] These data suggest that structural damage can already be found very early in the disease.

Longitudinal studies that allow a proper evaluation of change from nr-axSpA to r-axSpA are scarce: Sampaio-Barros *et al* found a 10% progression rate over 2 years in one study[18] and a 24% progression rate over 10 years in another study.[19] However, only the researchers of the GESPIC cohort realised that a proper



A = Switch from nr to r-axSpA according to the mNY criteria (worsened) *minus* switch from r to nr-axSpA (N=416)
B = Change in at least one grade in at least one SIJ (N=408)
C = Change in at least one grade in at least one SIJ and a final (at year 5) absolute value of at least 2 in the worsened joint (worsened) *minus* change in at least one grade in at least 1 SIJ and a baseline (year 0) absolute value of at least 2 in the improved joint (N=408)

Figure 1 Changes in different binary SIJ-Plain X-ray outcome measures (completers' population). nr-axSpA, radiographic axial spondyloarthritis; r-axSpA, radiographic axial spondyloarthritis; SIJ, sacroiliac joint.

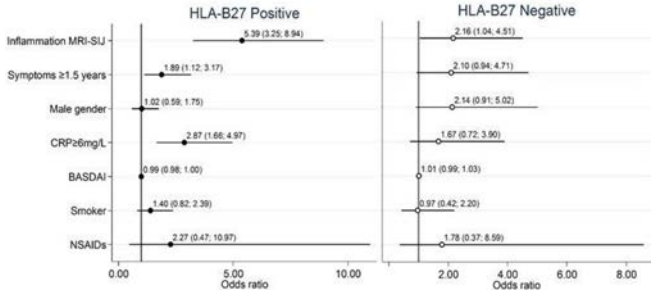


Figure 2 Effect of inflammation on MRI-SIJ on being mNY-positive after 5 years irrespective of baseline mNY status stratified according to the HLA-B27 status at baseline (1-level binomial multivariable GEE). Interaction between inflammation on MRI-SIJ and HLA-B27 at baseline: p=0.033. BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, creactive protein; GEE, generalised estimating equations; mNY, modified New York criteria; MRI-SIJ, magnetic resonance imaging of the sacroiliac joints; NSAIDs, non-steroidal anti-inflammatory drugs.

Table 2 Sensitivity analyses: effect of baseline MRI-SIJ inflammation on different SIJ radiographic progression definitions, irrespective of baseline mNY status and using different analytical approaches

	Main effect aOR (95% CI)	HLA-B27 positive aOR (95% CI)	HLA-B27negative aOR (95% CI)	p Value interaction
Outcome: mNY positive				
Logistic regression*	NA	9.26(4.32 to 19.86) (n=247)	3.79 (1.01 to 14.28) (n=143)	0.106
Logistic regression after MI†	6.64 (3.67 to 12.00) (n=557)	NA	NA	NS
1-level GEE‡	NA	5.39 (3.25 to 8.94) (n=248)	2.16 (1.04 to 4.51) (n=143)	0.033
2-level GEE (longitudinal)§	2.42 (1.01 to 5.78) (n=493)	NA	NA	NS
Outcome: 1-grade progression				
Logistic regression*	2.33 (1.21 to 4.49) (n=373)	NA	NA	NS
Logistic regression after MI†	2.35 (1.13 to 4.86) (n=557)	NA	NA	NS
1-level GEE‡	1.74 (1.05 to 2.88) (n=381)	NA	NA	NS
2-level GEE (longitudinal)§	1.90 (1.16 to 3.13) (n=486)	NA	NA	NS
Outcome: 1-grade progression + follow-up grade≥2				
Logistic regression*	3.45 (1.65 to 7.23) (n=373)	NA	NA	NS
Logistic regression after MI†	3.47 (1.60 to 7.54) (n=557)	NA	NA	NS
1-level GEE‡	1.82 (1.02 to 3.27) (n=381)	NA	NA	NS
2-level GEE (longitudinal)§	1.87 (1.04 to 3.36) (n=486)	NA	NA	NS

*Association between baseline MRI-SIJ inflammation and the X-SIJ score at year 5 with both variables according to the '2 out of 3' definition; N=patients with X-SIJ score available at year 5 and complete data on all covariates at baseline.
†Association between baseline MRI-SIJ inflammation and the X-SIJ score at year 5 both variables according to the '2 out of 3' definition, after multiple imputation; N= patients with X-SIJ available at baseline and in at least one postbaseline visit and complete data on all covariates at baseline.
‡Association between baseline MRI-SIJ inflammation and the X-SIJ score at year 5 incorporating measurements from all readers at baseline for MRI-SIJ and year 5 for the X-SIJ score and taking into account the within-reader correlation; N=patients with at least one baseline MRI-SIJ/5-year X-SIJ pair (ie, at the same time points available) and complete data on all covariates at baseline.
§Longitudinal association between MRI-SIJ inflammation and X-SIJ score (all measurements from all readers for both modalities) over the 5-year follow-up with time-lagged models and first-order autoregression, taking into account the within-reader and within-patient correlation for the repeated measurements; N=patients with at least one X-SIJ/MRI-SIJ pair and complete data on all covariates for the available pairs.
aOR, adjusted OR (adjusted for: symptom duration, gender, CRP, BASDAI, smoking status, treatment with NSAIDs and treatment with TNFi for longitudinal models); BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C reactive protein; GEE, generalised estimating equations; MI, multiple imputation; mNY, modified New York criteria; MRI-SIJ, MRI of the sacroiliac joints; NSAIDs, non-steroidal anti-inflammatory drugs; NA, not applicable—the main effect of MRI-SIJ inflammation on the different outcomes is only shown if the interaction with HLA-B27 is not significant (p≥0.15); NS, not significant; otherwise the effect of MRI-SIJ in each strata of HLA-B27 is shown; TNFi, tumour necrosis factor inhibitors.

progression estimate should aggregate worsening as well as improvement and reported progression in 9% after 2 years.[17]

The mNY criteria that quantify radiographic damage in SIJ have been proposed several decades ago for classifying a particular patient at a particular point of time. These inherently binary criteria (mNY+ or mNY−) were not intended to evaluate the natural course of the disease. Adaptations thereof may be more sensitive to change and simpler to interpret: our continuous score modification (a score from 0 to 8 based on the ordinal scale of mNY grading) is more sensitive but harder to interpret to the data analyst and the clinician. The statistician will worry about the handling of a semiquantitative variable as if it were a continuous one and will argue the seemingly similar distance between different grades. Moreover, a continuous score is simply the sum of the scores obtained in two SIJs, as if they were independent. A simpler means to express progression to the clinician is to define progression as a change of at least 1 grade in at least one SIJ. This proposal has been used for the first time by the GESPIC researchers.[16] Since we felt that a change between grade 0 and grade 1 (and vice versa) is not clinically relevant, we proposed a third definition by ignoring a change from 0 to 1.[3] Our study has confirmed that the sensitivity to change of this adjusted definition is better than the one based on the mNY criteria.

The main weakness of these X-SIJ-based definitions is likely the poor interobserver reliability: the assessment of radiographic damage in the SIJ according to the binary mNY criteria is particularly susceptible to measurement error.[20] While trained central readers have shown better reliability than single (local) readers, a combined-score by our three central readers ('2 out of 3' score) is still fallible in terms of measurement error, as is suggested by the finding of 'improvement' of SIJ damage under fully blinded conditions in a significant proportion of patients.

This means that measurement error (ie, scoring variability) must be taken into account when analysing X-SIJ progression. We have addressed this in two ways: first, our analysis was

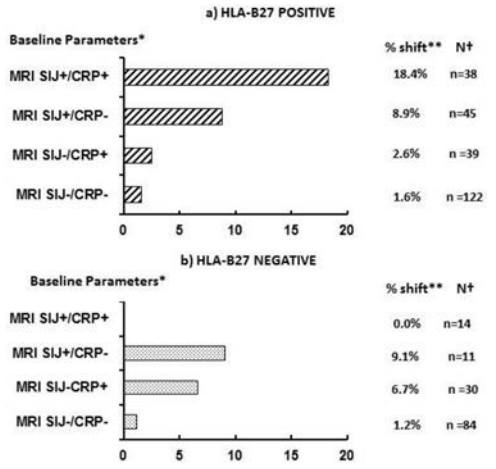


Figure 3 Net progression from nr-axSpA to r-axSpA according to baseline objective inflammatory markers and stratified on HLA-B27 ststus. BMO, bone marrow oedema; CRP, C reactive protein; MRI-SIJ, MRI of the sacroiliac joints; nr-axSpA, non-radiographic axial spondyloarthritis; r-axSpA, radiographic axial spondyloarthritis.

assumption free. We allowed 'positive change' as well as 'negative change' to occur without labelling this as 'true progression' or 'noise'. We analysed to what extent 5-year SIJ structural damage was driven by baseline inflammation on MRI-SIJ, and we could confirm a positive association: more MRI-inflammation at baseline leads to a higher 5-year SIJ score. In addition, we have used an analytical approach that most efficiently captures all the available information in the model, which adds to precision. In fact, our main analysis (the 1-level GEE) was more precise (narrower CI) than the 'traditional' logistic regression.

The other weakness of the X-SIJ outcome measures is the lack of information concerning their external validity and in particular the lack of information related to the impact of the changes in X-SIJ on the patient's functional disability. In this regard, syndesmophyte development at the spine level might be more relevant.

This cohort study in early axSpA reiterates the importance of BMO on MRI-SIJ as a predisposing factor for developing radiographic sacroiliitis 5 years later.[3 20] Of note, HLA-B27 was an effect modifier: patients carrying this genetic (risk) marker had a larger effect of MRI inflammation on radiographic damage than those not carrying this marker. This disparate effect suggests HLA-B27 is a critical factor for the severity of axSpA.[21 22]

Our data suggest that a proper risk estimation in individual patients is within our scope: an nr-axSpA patient that is HLA-B27-negative has a normal CRP and a negative MRI-SIJ has a likelihood of only 1.2% to progress to r-axSpA. In contrast, this likelihood is 18.4%; if the patient is HLAB27-positive, the CRP is increased and the MRI-SIJ shows BMO.

Further studies are required to better estimate the X-SIJ progression in axSpA and to better understand the role of inflammation on this progression.

Acknowledgements The DESIR cohort was sponsored by the Département de la Recherche Clinique et du Développement de l'Assistance Publique–Hôpitaux de Paris. This study is conducted under the umbrella of the French Society of Rheumatology and INSERM (Institut National de la Santé et de la Recherche Médicale). The database management is performed within the department of epidemiology and biostatistics (Professor Paul Landais, D.I.M., Nîmes, France). An unrestricted grant from Pfizer was allocated for the 10 years of the follow-up of the recruited patients. The authors would like to thank the different regional participating centres: Pr Maxime Dougados (Paris – Cochin B), Pr André Kahan (Paris - Cochin A), Pr Olivier Meyer (Paris - Bichat), Pr Pierre Bourgeois (Paris - La Pitié Salpetrière), Pr Francis Berenbaum (Paris - Saint Antoine), Pr Pascal Claudepierre (Créteil), Pr Maxime Breban (Boulogne Billancourt), Dr Bernadette Saint-Marcoux (Aulnay-sous-Bois), Pr Philippe Goupille (Tours), Pr Jean-Francis Maillefert (Dijon), Dr Xavier Puéchal, Dr Emmanuel Dernis (Le Mans), Pr Daniel Wendling (Besançon), Pr Bernard Combe (Montpellier), Pr Liana Euller-Ziegler (Nice), Pr Philippe Orcel, Dr Pascal Richette (Paris - Lariboisière), Pr Pierre Lafforgue (Marseille), Dr Patrick Boumier (Amiens), Pr Jean-Michel Ristori, Pr Martin Soubrier (Clermont-Ferrand), Dr Nadia Mehsen (Bordeaux), Pr Damien Loeuille (Nancy), Pr René-Marc Flipo (Lille), Pr Alain Saraux (Brest), Pr Corinne Miceli (Le Kremlin Bicêtre), Pr Alain Cantagrel (Toulouse), Pr Olivier Vittecoq (Rouen). The authors would also like to thank URC-CIC Paris Centre for the coordination and monitoring of the study.

Contributors All authors contributed and finally approved the current manuscript.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Comitte de Protection des Personnes Ile de France III.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Dougados M, Baeten D. Spondyloarthritis. *Lancet* 2011;377:2127–37.
2 Sieper J, Braun J, Dougados M, *et al*. Axial spondyloarthritis. *Nat Rev Dis Primers* 2015;1:15013.
3 Dougados M, Demattei C, van den Berg R, *et al*. Rate and Predisposing Factors for Sacroiliac Joint Radiographic Progression After a Two-Year Follow-up Period in Recent-Onset Spondyloarthritis. *Arthritis Rheumatol* 2016;68:1904–13.
4 Dougados M, Etcheto A, Molto A, *et al*. DESIR cohort. Clinical presentation of patients suffering from recent onset chronic inflammatory back pain suggestive of spondyloarthritis: The DESIR cohort. *Joint Bone Spine* 2015;82:345–51.
5 Calin A, Porta J, Fries JF, *et al*. Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977;237:2613–4.
6 Rudwaleit M, Metter A, Listing J, *et al*. Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. *Arthritis Rheum* 2006;54:569–78.
7 Garrett S, Jenkinson T, Kennedy LG, *et al*. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. *J Rheumatol* 1994;21:2286–91.
8 Calin A, Garrett S, Whitelock H, *et al*. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *J Rheumatol* 1994;21:2281–5.
9 Dougados M, Simon P, Braun J, *et al*. ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. *Ann Rheum Dis* 2011;70:249–51.
10 van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361–8.
11 Rudwaleit M, Jurik AG, Hermann KG, *et al*. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann Rheum Dis* 2009;68:1520–7.
12 Mau W, Zeidler H, Mau R, *et al*. Clinical features and prognosis of patients with possible ankylosing spondylitis. Results of a 10-year followup. *J Rheumatol* 1988;15:1109–14.
13 Heuft-Dorenbosch L, Weijers R, Landewé R, *et al*. Magnetic resonance imaging changes of sacroiliac joints in patients with recent-onset inflammatory back pain: inter-reader reliability and prevalence of abnormalities. *Arthritis Res Ther* 2006;8:R11.
14 Ramiro S, van der Heijde D, van Tubergen A, *et al*. Higher disease activity leads to more structural damage in the spine in ankylosing spondylitis: 12-year longitudinal data from the OASIS cohort. *Ann Rheum Dis* 2014;73:1455–61.
15 Poddubnyy D, Brandt H, Vahldiek J, *et al*. The frequency of non-radiographic axial spondyloarthritis in relation to symptom duration in patients referred because of chronic back pain: results from the Berlin early spondyloarthritis clinic. *Ann Rheum Dis* 2012;71:1998–2001.
16 Said-Nahal R, Miceli-Richard C, Berthelot JM, *et al*. The familial form of spondylarthropathy: a clinical study of 115 multiplex families. Groupe Français d'Etude Génétique des Spondylarthropathies. *Arthritis Rheum* 2000;43:1356–65.
17 Poddubnyy D, Rudwaleit M, Haibel H, *et al*. Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. *Ann Rheum Dis* 2011;70:1369–74.

18 Sampaio-Barros PD, Conde RA, Donadi EA, *et al*. Undifferentiated spondyloarthropathies in Brazilians: importance of HLA-B27 and the B7-CREG alleles in characterization and disease progression. *J Rheumatol* 2003;30:2632–7.
19 Sampaio-Barros PD, Bortoluzzo AB, Conde RA, *et al*. Undifferentiated spondyloarthritis: a longterm followup. *J Rheumatol* 2010;37:1195–9.
20 van den Berg R, Lenczner G, Feydy A, *et al*. Agreement between clinical practice and trained central reading in reading of sacroiliac joints on plain pelvic radiographs. Results from the DESIR cohort. *Arthritis Rheumatol* 2014;66:2403–11.
21 Chung HY, Machado P, van der Heijde D, *et al*. HLA-B27 positive patients differ from HLA-B27 negative patients in clinical presentation and imaging: results from the DESIR cohort of patients with recent onset axial spondyloarthritis. *Ann Rheum Dis* 2011;70:1930–6.
22 Bennett AN, McGonagle D, O'Connor P, *et al*. Severity of baseline magnetic resonance imaging-evident sacroiliitis and HLA-B27 status in early inflammatory back pain predict radiographically evident ankylosing spondylitis at eight years. *Arthritis Rheum* 2008;58:3413–8.

OPEN ACCESS

EXTENDED REPORT

Predicting and managing primary and secondary non-response to rituximab using B-cell biomarkers in systemic lupus erythematosus

Md Yuzaiful Md Yusof,[1,2] Daniel Shaw,[1] Yasser M El-Sherbiny,[1,2,3] Emma Dunn,[4] Andy C Rawstron,[5] Paul Emery,[1,2] Edward M Vital[1,2]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211191).

[1]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Chapel Allerton Hospital, Chapeltown Road, Leeds, UK
[2]NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[3]Department of Clinical Pathology, Faculty of Medicine, Mansoura University, Egypt
[4]Department of Nephrology, St James' University Hospital, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[5]Haematological Malignancy Diagnostic Service, Leeds Teaching Hospitals NHS Trust, Leeds, UK

Correspondence to
Dr Edward M Vital, Chapel Allerton Hospital, Leeds LS7 4SA, UK;
e.m.j.vital@leeds.ac.uk, edvital@gmail.com

PE and EMV contributed equally.

Received 23 January 2017
Revised 27 April 2017
Accepted 6 June 2017
Published Online First 6 July 2017

To cite: Md Yusof MY, Shaw D, El-Sherbiny YM, *et al*. *Ann Rheum Dis* 2017;**76**:1829–1836.

ABSTRACT

Objective To assess factors associated with primary and secondary non-response to rituximab in systemic lupus erythematosus (SLE) and evaluate management of secondary non-depletion non-response (2NDNR).

Methods 125 patients with SLE treated with rituximab over 12 years were studied prospectively. A major clinical response was defined as improvement of all active British Isles Lupus Assessment Group (BILAG)-2004 domains to grade C/better and no A/B flare. Partial responders were defined by one persistent BILAG B. B-cell subsets were measured using highly sensitive flow cytometry. Patients with 2NDNR, defined by infusion reaction and defective depletion, were treated with ocrelizumab or ofatumumab.

Results 117 patients had evaluable data. In cycle 1 (C1), 96/117 (82%) achieved BILAG response (major=50%, partial=32%). In multivariable analysis, younger age (OR 0.97, 95% CI 0.94 to 1.00) and B-cell depletion at 6 weeks (OR 3.22, 95% CI 1.24 to 8.33) increased the odds of major response. Complete depletion was predicted by normal complement and lower pre-rituximab plasmablasts and was not associated with increased serious infection post-rituximab. Seventy-seven (with data on 72) C1 responders were retreated on clinical relapse. Of these, 61/72 (85%) responded in cycle 2 (C2). Of the 11 C2 non-responders, nine met 2NDNR criteria (incidence=12%) and tested positive for anti-rituximab antibodies. Lack of concomitant immunosuppressant and higher pre-rituximab plasmablasts predicted 2NDNR. Five were switched to ocrelizumab/ofatumumab, and all depleted and responded.

Conclusion Treatment with anti-CD20 agents can be guided by B-cell monitoring and should aim to achieve complete depletion. 2NDNR is associated with anti-rituximab antibodies, and switching to humanised agents restores depletion and response. In SLE, alternative anti-CD20 antibodies may be more consistently effective.

INTRODUCTION

Rituximab, a chimeric anti-CD20 monoclonal antibody (mAb) remains an important treatment option for moderate to severe systemic lupus erythematosus (SLE). A high degree of efficacy of rituximab across a range of lupus manifestations has been reported in open-label studies from single-centre series,[1–3] multicentre registries[4–6] and a systematic review of off-label use.[7] Despite the success of these series, two phase III randomised placebo-controlled trials in non-renal lupus[8] and renal lupus[9] failed to meet their primary end-points. The discrepancy between the randomised trials and real-world evidence has been attributed to aspects of trial design including choice of end-points, the use of an active comparator, inclusion criteria and low statistical power.[10]

Nevertheless, there are also mechanistic reasons for the failure of rituximab in clinical trials in SLE. B-cell killing by rituximab appeared less efficient in SLE than rheumatoid arthritis (RA)[11] due to internalisation through interaction with FcγRIIb resulting in reduced effector activity[12] and pathogenic lupus autoantibodies that were produced by long-lived plasma cells.[13 14] Using highly sensitive flow cytometry (HSFC), a protocol that was optimised for the detection of plasmablasts, we discovered that the depth of B-cell depletion predicted response in RA[15] and SLE.[2] Similar studies as well as identifying other clinical predictors of response to rituximab in SLE are needed to optimise its use and to help design trials of alternative B-cell depleting strategies.

B-cell depletion therapy with rituximab is transient. Some patients with initial good response experience relapse after B-cell repopulation (although with a variable interval). In our published discovery cohort, we showed a bimodal pattern of relapse. Earlier relapse requiring rituximab retreatment was predicted by a plasmablast count of $>0.0008\times10^9$/L at 6 months (the time of initial clinical response).[2] Patients with lower plasmablasts at 6 months had sustained response without retreatment. Validation of this as a biomarker is therefore needed to determine whether HSFC can be used in clinical practice to guide retreatment decisions.

Repeat treatment with rituximab is effective.[1] However, we observed cases of patients with SLE who had previously depleted and responded well to rituximab but subsequently developed (1) a severe infusion reaction >24 hours during the second infusion of a cycle, (2) failure to deplete CD20+ (naïve and memory) B-cells and (3) clinical non-response during repeat cycles. We called this phenomenon secondary non-depletion and non-response (2NDNR), which was suggestive of immunogenicity to rituximab and could be overcome by alternative anti-CD20 mAbs, particularly humanised. Therefore, the aims of the study were to assess factors predicting primary and secondary non-response to rituximab in SLE including validation of

B-cell depletion and to evaluate management of 2NDNR using alternative anti-CD20 agents.

METHODS

Patients and design

A prospective observational study was conducted of all patients with moderate to severe SLE who were treated with rituximab in Leeds between January 2004 and July 2016. Inclusion criteria included (1) adults (>16 years old); (2) fulfilling the revised 1997 American College of Rheumatology classification for SLE[16] and (3) at least 6 months follow-up post-rituximab.

Treatment protocol

All patients received a first cycle of therapy consisting of 100 mg of methylprednisolone and 1000 mg of rituximab given intravenously on days 1 and 14. Further cycles of the same regimen were repeated on clinical relapse (defined below).

Of those who met 2NDNR criteria, their treatment was switched from rituximab to humanised anti-CD20 mAbs either by using (1) 2×1000 mg ocrelizumab (compassionate use from Roche UK) or (2) 2×700 mg ofatumumab (individual funding request to NHS England).

Clinical data and outcomes

Disease activity was assessed using the British Isles Lupus Assessment Group (BILAG-2004)[17] at baseline and every 3 months thereafter. Clinical responses at 6 months were determined as following: (1) major clinical response=improvement of all domains rated A/B to grade C/better and no A/B flare between baseline and 6 months; (2) partial clinical response=maximum of 1 domain with a persistent grade B with improvement in all other domains and no A or B flare and (3) non-response=those not meeting the criteria for major or partial clinical response. Relapse was defined as a new grade A or recurrence of ≥1 grade B following either major/partial clinical response at 6 months. Global BILAG score was calculated as follows: grade A=12, grade B=8, grade C=1 and grades D and E=0.[18]

Laboratory assessments

Peripheral blood B-cell subsets (naïve, memory B-cells and plasmablasts) were measured using HSFC as previously described[15] at baseline, 6 months and every 6 months without knowledge of clinical status other than time since rituximab. Complete B-cell depletion was defined as counts $<0.0001\times10^9$/L and repopulation as $\geq0.0001\times10^9$/L.

Anti-dsDNA antibody titres were measured by ELISA until July 2012 and Bioplex 2200 Immunoassay (after July 2012). Complement levels (C3 and C4) and total serum immunoglobulin titres were measured by nephelometry.

Anti-rituximab antibodies were tested on a subset of patients with 2NDNR using the Promonitor® Anti-Rituximab ELISA according to the manufacturer's instructions and compared these concentrations to those with continued response to rituximab. A positive test (as determined by the manufacturer) was concentration >140 AU/mL.

Safety

Serious infections were recorded irrespective of suspected association with SLE and/or therapy. These were infections that resulted in hospitalisation for >24 hours or required intravenous antibiotics. Details about other safety assessment can be found in online supplementary files.

Statistical analysis

Descriptive statistics were summarised using mean with SD or median with IQR for continuous variables and proportion for categorical variables. Multiple imputation was used for missing data. Multivariable analyses were performed using logistic regression after checking for multicollinearity. The significance of the association between categorical variables was tested by Fisher's exact test, while for continuous variables using Mann-Whitney U test. Receiver operator curves (ROCs) were used to measure sensitivity and specificity of optimal thresholds for investigations predicting time-to-clinical relapse.

All statistical analysis was performed using Stata V.13.1 and Graph Pad Prism V.6.01 for Windows.

RESULTS

Patient characteristics

Of 125 patients with SLE who were treated with rituximab at our unit, 117 patients with evaluable data at 6 months were studied. Baseline characteristics are described in table 1. One hundred and twelve (96%) had refractory and active disease as defined by BILAG ≥1A score and/or ≥2B scores. The remaining five had BILAG B in one domain only but was refractory to other conventional therapies as well as on maintenance with oral prednisolone ≥10 mg daily. Total follow-up was 492 patient-years.

Treatment characteristics

Three hundred and eighteen cycles of rituximab were administered. Median (range) duration of response in rituximab responders for cycles 1–4 (C1–4) were 52 (26–423), 52 (26—299), 57 (27–184) and 50 (29–173) weeks, respectively.

Concomitant cyclophosphamide was used in five patients who presented with life-threatening flare.

Clinical and immunological response to first cycle rituximab

In C1, there was a good overall clinical response to rituximab. Fifty-eight (50%) patients had major clinical response, 38 (32%) partial clinical response and 21 (18%) were non-responders. The median global BILAG scores had reduced from 21 (IQR 14–27) pre-rituximab to 8 (IQR 1–10) at 6 months; p<0.001.

Responses in individual BILAG domains are shown in figure 1A. Although majority of domains improved, responses were more variable in the mucocutaneous and haematological domains. Mucocutaneous responses to rituximab have been described in detail previously.[19] These long-term data showed a more consistent major response in lupus erythematosus non-specific lesions and oral ulcers, while non-response in chronic cutaneous lupus erythematosus (CCLE) (CCLE vs other lupus-specific lesions; p=0.022).

The median serum anti-dsDNA titre had reduced from 109 (IQR 16–300) IU/mL pre-rituximab to 32 (IQR 7–116) IU/mL at 6 months; p<0.001. Of 46 patients with low complement (C3 and/or C4) levels pre-rituximab, levels had normalised in 25/46 (54%) at 6 months.

Predictors of major clinical response to first cycle rituximab

Only B-cell depletion at 6 weeks increased the odds of BILAG response (major/partial) in multivariable analysis; adjusted imputed OR 13.93, 95% CI 3.11 to 62.37; p=0.001 (online supplementary table S2).

As there was a high degree of response to rituximab in this cohort, we analysed predictors for major clinical response separately in order to identify patients who would respond best to therapy. In imputed univariable analysis, only younger age was

Table 1 Baseline characteristics of the 117 patients with SLE treated with rituximab

Age at first RTX infusion, median (IQR) years	39 (26–52)	
No. female patient (%)	109 (93)	
Ethnicity, N (%)		
Caucasian	80 (68)	
Afro-Caribbean	11 (10)	
South Asian	20 (17)	
Other	6 (5)	
SLE disease duration at first RTX, median (IQR) years	6 (2–11)	
Positive ANA at diagnosis, N (%)	117 (100)	
Antibody status at first RTX infusion, N (%) Positive	108 (92)	
anti-dsDNA	56 (48)	
Anti-Ro	57 (49)	
Anti-La	18 (15)	
Anti-Smith	15 (13)	
Anti-Chromatin	19 (16)	
Anti-RNP	23 (20)	
Anti-Ribosomal P	6 (5)	
Anti-Cardiolipin/anti-B2-glycoprotein	14 (12)	
Prior CYC therapy, N (%)	63 (54)	
Cumulative dose of CYC, mean ± SD gram	6.6 ± 4.2	
Number of prior immunosuppressant failure (including CYC but excluding glucocorticoid), median (range)	3 (0–9)	
Concomitant antimalarials, N (%)	88 (75)	
Concomitant immunosuppressant, N (%)		
Azathioprine	19 (16)	
Methotrexate	16 (14)	
Mycophenolate Mofetil	39 (33)	
Prednisolone dose at first RTX infusion, median (IQR) mg	10 (3–20)	
ESR at first RTX infusion, median (IQR) mm/hour	29 (15–57)	
BILAG index score at baseline, N (%)		
≥1 A score	96 (82)	
No A score but ≥2 B scores	16 (14)	
BILAG domains at baseline, N (%)	Grade A	Grade B
General	9 (8)	12 (10)
Mucocutaneous	23 (20)	32 (27)
Neurological	17 (15)	17 (15)
Musculoskeletal	30 (26)	24 (20)
Cardiorespiratory	6 (5)	13 (11)
Gastrointestinal	6 (5)	0 (0)
Ophthalmic	0 (0)	0 (0)
Renal	34 (29)	0 (0)
Haematology	11 (9)	12 (10)
Global BILAG score, median (IQR)	21 (14–27)	
SLEDAI-2K score, median (IQR)	10 (6–14)	
SLICC Damage Index, median (IQR)	0 (0–1)	

ANA, antinuclear antibody; BILAG, British Isles Lupus Assessment Group; CYC, cyclophosphamide; dsDNA, double-stranded DNA; ESR, erythrocyte sedimentation rate; RNP, ribonucleic protein; RTX, rituximab; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index 2000; SLICC, Systemic Lupus International Collaborating Clinics (SLICC).

associated with major response to rituximab (OR 0.97, 95% CI 0.95 to 0.99; p=0.031). While in imputed multivariable model, younger age (OR 0.97, 95% CI 0.94 to 1.00; p=0.045) and B-cell depletion at 6 weeks post-rituximab (OR 3.22, 95% CI 1.24 to 8.33; p=0.016) increased the odds of major response to rituximab (table 2).

Validation of association between complete B-cell depletion and clinical response

The published discovery cohort included 37 patients with SLE.[2] In this validation cohort, 67 subsequent and consecutive patients (with B-cell data available) were analysed. Similar to the discovery cohort, higher response rate was achieved in complete depletion compared with incomplete depletion groups (93% vs 68%; p=0.011) in this validation cohort (figure 1B).

While there was no difference at baseline, patients with complete B-cell depletion had significantly lower anti-dsDNA antibody titres at 14 weeks (p=0.030) and 26 weeks (p=0.041) versus those with incomplete depletion. In the former, C3 and C4 levels were not different at 14 weeks (p=0.064 and p=0.148, respectively) but were higher at 26 weeks (p=0.020 and p=0.022, respectively) compared with the latter group. There was no difference in anti-ENA antibodies between the two groups at 14 and 26 weeks; all p>0.10.

Predictors for complete B-cell depletion to first cycle rituximab

Data for B-cell subsets were available for 104 (89%) patients. In imputed univariable analysis, higher anti-dsDNA titre (OR 1.00, 95% CI 0.99 to 1.00; p=0.038), normal complement levels (OR 0.41, 95% CI 0.18 to 0.91; p=0.028) and lower pre-rituximab plasmablasts (OR 0.88, 95% CI 0.80 to 0.98; p=0.015) were associated with complete B-cell depletion. While in imputed multivariable model, only normal complement levels (OR 0.29, 95% CI 0.09 to 0.90; p=0.032) and lower pre-rituximab plasmablasts (OR 0.86, 95% CI 0.78 to 0.96; p=0.007) predicted complete B-cell depletion post-rituximab (online supplementary table S4).

B-cell depletion and associated serious infection

As most of the serious infection episodes occurred in C1 and C2 (n=23 in 15 patients), we analysed the association between complete B-cell depletion and serious infection. After two cycles, there were no difference in the serious infection rates between complete and incomplete depletion groups (8/98 (8.2%) and 7/73 (9.6%), respectively; p=0.789).

Plasmablast repopulation as a biomarker of relapse

At 6 months, B-cells were detectable in 81% of the C1 responders. This time-point preceded all relapses. As the median of duration of response was 52 weeks, we divided the patients in this validation cohort (n=25 with B-cells data available) into two groups: (1) earlier relapse (≤12 months from first rituximab) and (2) later relapse (>12 months). A 12-month relapse time is clinically significant as it indicates that a 6-monthly retreatment may not be necessarily needed in these patients. Similar to the discovery cohort, the ROC indicated that a plasmablast count of $>0.0008\times10^9$/L at 6 months yielded 73% (95% CI 45% to 92%) sensitivity and 90% (95% CI 56% to 99%) specificity in predicting earlier relapse; area under the curve of 0.86 (online supplementary figure S1).

Of the patients with plasmablasts $>0.0008\times10^9$/L at 6 months, relapse rates within the next 26 and 52 weeks were 90% and 100%, respectively. While of the patients with plasmablasts $\leq0.0008\times10^9$/L at 6 months, relapse rates within

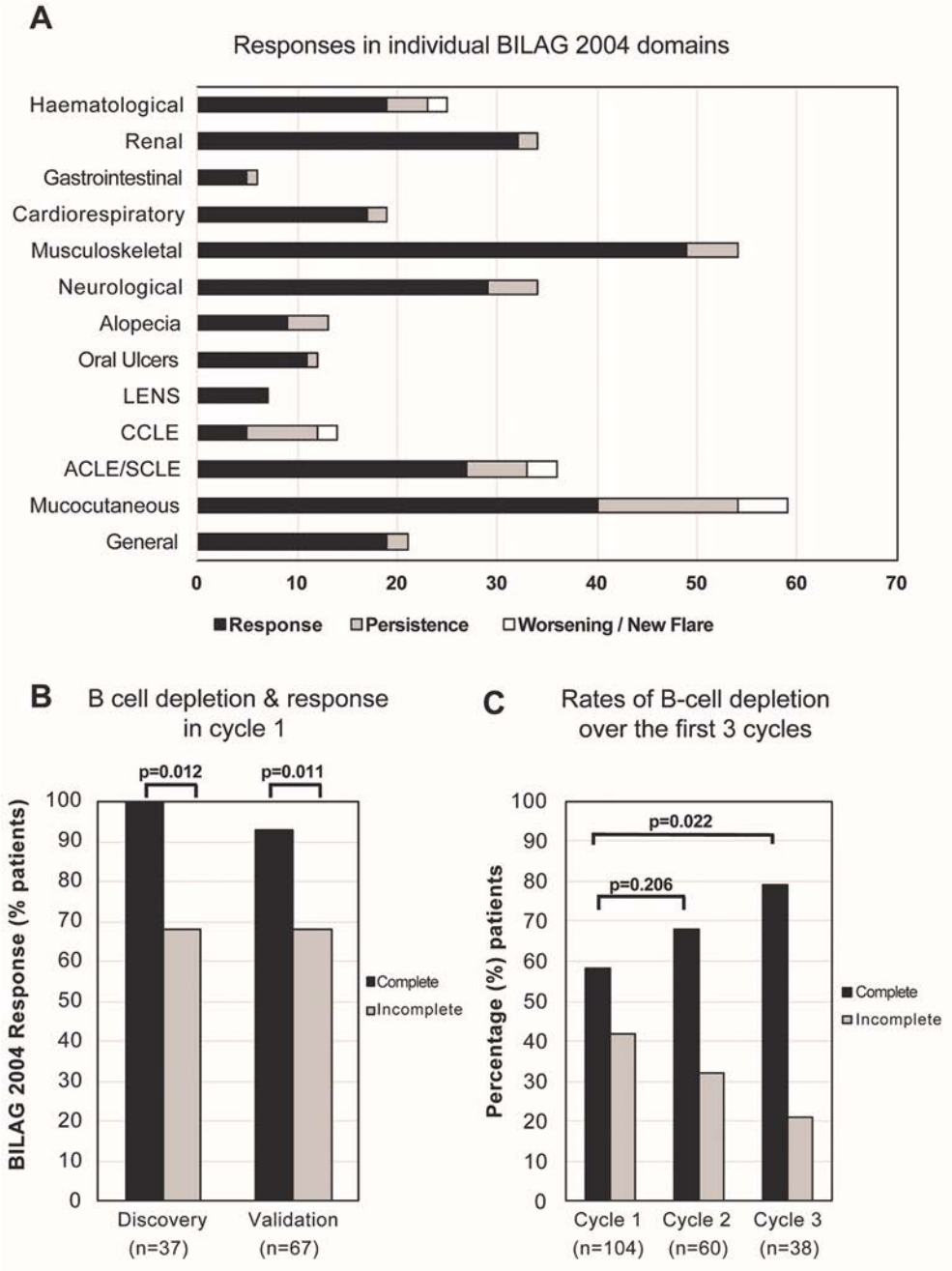


Figure 1 BILAG response and B-cell depletion following rituximab. (A) Majority of the individual domain improved post-rituximab although responses in the mucocutaneous and haematological domains were more varied. (B) Similar to the discovery cohort, a higher response rate was achieved in complete depletion compared with incomplete depletion groups; 93% versus 68%; p=0.011 in the validation cohort. (C) There was an incremental increase in the rates of B-cell depletion over three cycles of rituximab. ACLE, acute cutaneous lupus erythematosus; BILAG: British Isles Lupus Assessment Group; CCLE, chronic cutaneous lupus erythematosus; LENS, lupus erythematosus non-specific lesions.

the next 26 and 52 weeks were 33% and 73%, respectively (figure 2A).

There were no differences in anti-dsDNA titres, total BILAG score and memory B-cells at 6 months between the earlier versus later relapse groups, p=0.475, p=0.985 and p=0.414, respectively.

Retreatment of first cycle non-responders

In RA, we showed that retreatment of initial non-responders with incomplete B-cell depletion led to improved response rate in C2.[20] Of the 21 patients who were C1 non-responders, nine were retreated with rituximab. The domains that persisted at grade A/B in C1 were mucocutaneous (n=4), musculoskeletal (n=3), renal (n=2) and haematology (n=3). After retreatment, none of these patients responded. Additionally, four patients had clinical features that were suggestive of immunogenicity.

Retreatment of first cycle responders

Of the 96 patients who were C1 responders, 77 (with complete data on 72) were retreated on clinical relapse. Of these, 61/72 (85%) responded in C2 (figure 3). Numerically higher rate of B-cell depletion was achieved in C2 compared with C1 (68% versus 58%, respectively; p=0.206) and depletion improved

Table 2 Multivariable analysis for predictors of major clinical response to first cycle rituximab

	No response/ partial response n=59	Major clinical response n=58	Univariable OR (95% CI), p value (with multiple imputation)	Multivariable OR (95% CI), p value (with multiple imputation)
Age, mean (SD) years	43 (17)	37 (14)	**0.97 (0.95 to 0.99), p=0.031 per year**	**0.97 (0.94 to 1.00), p=0.045**
White, N (%)	43 (73)	37 (64)	1.53 (0.70 to 3.34), p=0.292	0.92 (0.34 to 2.47), p=0.870
Anti-dsDNA titres, mean (SD) IU/mL	147 (230)	142 (230)	1.00 (0.99 to 1.00), p=0.879 per unit	1.00 (0.99 to 1.00), p=0.632
Anti-ENA positivity, N (%)	40 (68)	38 (66)	0.91 (0.42 to 1.99), p=0.812	0.90 (0.37 to 2.22), p=0.821
Low C3 and/or C4 titres, N (%)	25 (42)	24 (41)	0.97 (0.46 to 2.04), p=0.937	1.14 (0.41 to 3.13), p=0.801
ESR, mean (SD) mm/hour*	40 (32)	41 (36)	1.00 (0.99 to 1.01), p=0.827 per unit	–
Concomitant S, N (%)†	41 (69)	35 (60)	0.67 (0.31 to 1.43), p=0.301	0.43 (0.17 to 1.09), p=0.075
Daily prednisolone dose, mean (SD) mg	13 (11)	16 (14)	1.02 (0.99 to 1.05), p=0.207 per mg	1.00 (0.97 to 1.04), p=0.713
Total BILAG score, mean (IQR)	21 (8)	24 (13)	1.03 (0.99 to 1.07), p=0.093 per point	1.02 (0.97 to 1.07), p=0.371
Total B-cell counts, mean (IQR)‡	101 (95)	138 (150)	1.00 (1.00 to 1.01), p=0.161 per unit	1.00 (1.00 to 1.01), p=0.137
B-cell depletion at 6 weeks postrituximab, N (%)	29 (49)	39 (68)	2.10 (0.95 to 4.62), p=0.065	**3.22 (1.24 to 8.33), p=0.016**

*As high collinearity was observed between ESR and total B-cell counts, only the latter was included in the multivariable analysis.
†Concomitant immunosuppressant was defined as either using methotrexate, azathioprine, mycophenolate mofetil and/or other disease modifying anti-rheumatic drugs but excluded anti-malarials.
‡count × 10^9 cells/L) for each subset multiplied by 1000 prior to analysis.
BILAG, British Isles Lupus Assessment Group; C3/C4, complement 3 or 4; dsDNA, double-stranded DNA; ENA, extract nuclear antigen; ESR, erythrocyte sedimentation rate; IS, immunosuppressant.

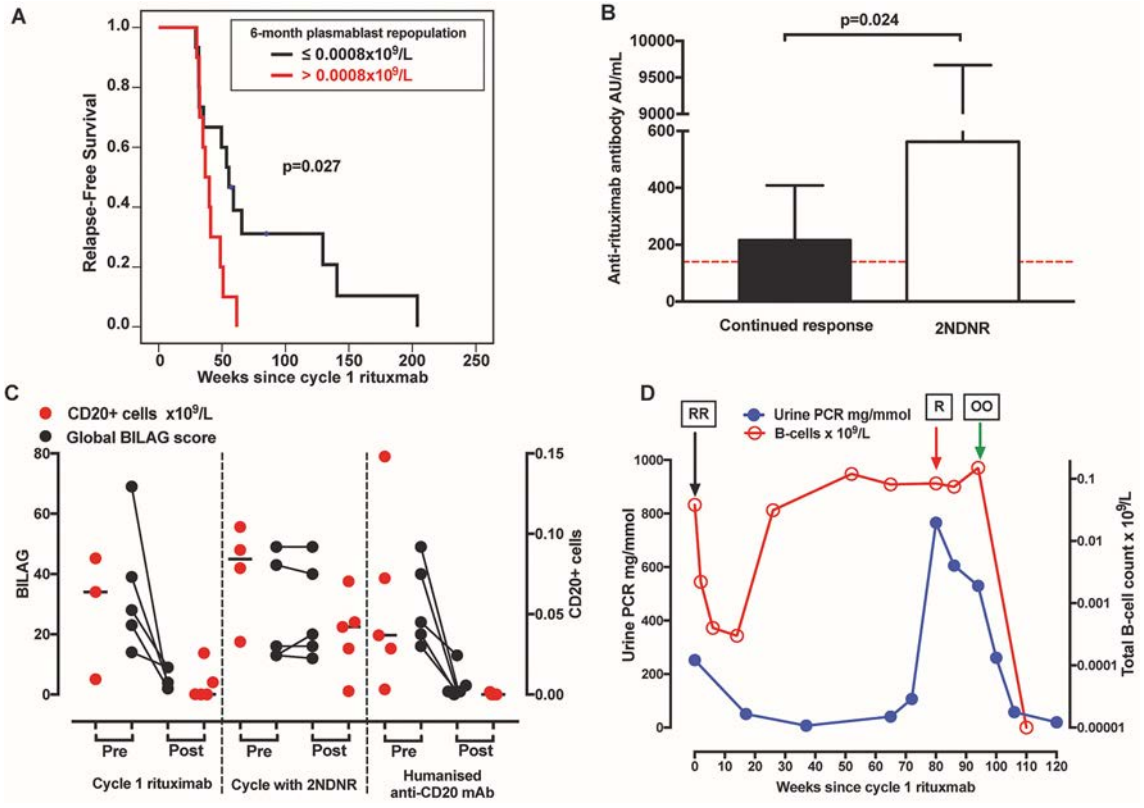


Figure 2 2NDNR to rituximab and efficacy of alternative humanised anti-CD20 antibodies. (A) In this validation cohort, detection of plasmablasts >0.0008×10^9/L at 6 months predicted earlier relapse. (B) The phenomenon 2NDNR was associated with anti-rituximab antibody. The dotted red line represents normal cut-off of the test. (C) The Global BILAG score and CD20+ B-cells are plotted for each patient. The black line in the CD20+ B-cells figure represents the median. (D) An example of a case where proteinuria was normalised following a switch to ocrelizumab. 'RR' represents 2x infusions of rituximab, 'R' represents a single infusion as the patient cannot not complete the second due to severe infusion reaction and 'OO' represents 2x infusions of ocrelizumab. The total B-cell counts were transformed to natural log. 2NDNR, secondary non-depletion non-response; BILAG, British Isles Lupus Assessment Group.

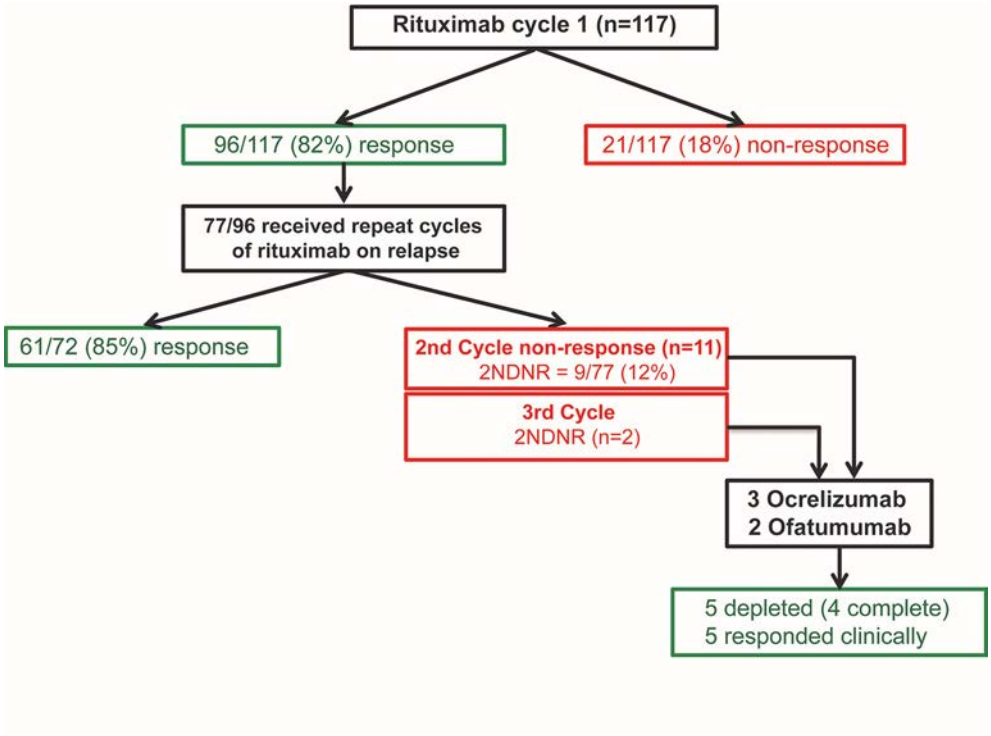


Figure 3 Efficacy of repeat cycles with rituximab in systemic lupus erythematosus. There was a high rate of initial clinical response to rituximab in this cohort, 96/117 (82%). Seventy-seven responders who had clinical relapse were retreated in C2. Of these, 61/72 (85%) continued to response in C2. Of the C2 non-responders, 9/11 met 2NDNR criteria. Five were switched to ocrelizumab/ofatumumab resulted in depletion and response in all. 2NDNR, secondary non-depletion and non-response; C1, cycle 1.

Table 3 Factors associated with secondary non-depletion non-response to rituximab (2NDNR)

Characteristics prior to rituximab retreatment	Continued to respond (n=61)	2NDNR (n=9)	p Value
Concomitant IS, N (%)	41 (67)	2 (22)	**0.023**
Prednisolone, median (IQR) mg	5 (0–10)	5 (0–17.5)	0.729
Duration of response, median (IQR) weeks	50 (36–107)	62 (52–164)	0.239
Total BILAG score, median (IQR)	16 (12–21)	24 (12–27)	0.209
Partial clinical response in cycle 1, N (%)	24 (39)	3 (33)	0.731
Naïve B-cells, median (IQR) 10^9 cells/L	0.0349 (0.0071–0.0735)	0.0620 (0.0101–0.0950)	0.296
Memory B-cells, median (IQR) x 10^9/L	0.0019 (0.0010–0.0047)	0.0090 (0.0054–0.0394)	0.175
Plasmablasts, median (IQR) x 10^9/L	0.0011 (0.0004–0.0036)	0.0086 (0.0052–0.0227)	**<0.001**

*NDNR, secondary non-depletion and non-response; IS, immunosuppressant.

over subsequent cycle, C3 versus C1 (79% vs 58% respectively; p=0.022) (figure 1C).

Twelve out of 38 patients who were C1 partial responders were retreated at 6 months. Of these, major clinical response was achieved in 10/12 (83%) in C2. One patient had worsening of arthritis, while another had 2NDNR in C2.

Of the 11 patients who were C2 non-responders, nine met 2NDNR criteria. Therefore, the incidence of 2NDNR in this cohort was 9/77 (12%). In C3, another two patients had 2NDNR.

Association of 2NDNR with antirituximab antibody

Post-rituximab sera for 5/9 patients with 2NDNR were tested for anti-rituximab antibodies. Of these, all 5/5 (100%) were tested positive. In contrast, of the 16 patients who were C2 responders, 9/16 (56%) were also tested positive for anti-rituximab antibodies. The median anti-rituximab levels were higher in the former, 562 (IQR 394–9670) AU/mL compared with the latter, 217 (IQR 0–409) AU/mL; p=0.024 (figure 2B).

Factors associated with 2NDNR

Risk factors for 2NDNR were lack of concomitant immunosuppressant (p=0.023) and higher pre-rituximab plasmablasts (p<0.001) (table 3). Concomitant corticosteroid dose, duration of response in C1, clinical response category in C1, pre-rituximab global BILAG score, pre-rituximab naïve and memory B-cells were not associated with 2NDNR; all p>0.10.

Efficacy of switching to humanised anti-CD20 antibodies

Following 2NDNR, treatment for five patients were switched to humanised anti-CD20 mAbs (3=ocrelizumab and 2=ofatumumab). Post-treatment, complete depletion of CD20+ cells were achieved in 4/5 patients, while the remaining one had substantially low counts (0.0016×10^9/L).

The median global BILAG scores had reduced from 24 (IQR 18–45) pre-treatment to 1 (IQR 0–8) post-treatment; p=0.008 (figure 2C). The individual BILAG response is shown in figure 2D and described in online supplementary table S5. One patient with

class IV-G (active with moderate scarring) who had progressed into end-stage renal failure was treated with ofatumumab, mainly for severe thrombocytopaenia with a view for renal transplantation preparation. Post-treatment, her platelet had normalised from 45×10^9/L (pre-treatment), renal parameters were stable and she successfully underwent live donor renal transplantation.

DISCUSSION

The clinical challenges for the use of rituximab in SLE include defining subgroups of patients likely to respond to the initial and subsequent cycles and optimal repeat treatment strategy. By capturing data of all patients with SLE who were treated with rituximab in this largest reported cohort, as well as long-term follow-up, this study offers insights into pragmatic use of rituximab and has implications for the future development of targeted therapies.

In this study, the only consistent predictor of any (and major) clinical response to rituximab is B-cell depletion (as measured using HSFC) at 6 weeks post-rituximab, which we have now validated in an independent cohort. This underlines the immunomodulatory action of rituximab in correcting autoimmune B-cell function and normalising autoantibody titres and complement levels without increasing the risk of severe infection. From treatment stratification perspective, our data support the rationale for B-cell monitoring during therapy. Thus, prior to rituximab, by assessing patients for low complement levels and higher plasmablasts, treatment modification can be employed to improve depletion, either by increasing the dose or adding an extra infusion, as we previously showed in RA.[21] At 6 weeks post-rituximab, complete depletion is a marker of good response to therapy. For those with incomplete depletion, close monitoring is required. At 6 months post-rituximab, repopulation of plasmablasts of $>0.0008\times10^9$/L increases the risk of clinical relapse within the following 6 months. Therefore, these patients can be considered for early retreatment in order to reduce the higher burden of B-cell numbers and enhance depletion in the subsequent cycle. Importantly, for those with plasmablasts of $\leq0.0008\times10^9$/L at 6 months, monitoring for clinical relapse would appear an acceptable strategy.

Regardless of response, about 12% subsequently developed 2NDNR in C2. This phenomenon is associated with rituximab anti-drug antibodies. However, measuring anti-rituximab antibody alone is not enough to identify patients as 2NDNR as over half of the patients who were tested positive responded in that particular cycle. Instead, clinical features, that is, severe infusion reaction and non-response and measuring B-cells, are more meaningful. Lack of concomitant oral immunosuppressant and higher pre-rituximab plasmablasts predicted 2NDNR. Oral immunosuppressant use was decided at physician discretion, but our data suggest they might have a role in preventing immunogenicity. The exact mechanism for the association with plasmablast number is unknown, but plasmablasts are markers for overall B-cell activation. Following initial depletion with rituximab, B-cell-activating factor levels increase and promote the formation of plasmablasts.[22] This early increase in plasmablasts enhances the formation of follicular T-helper cells, thus creating a positive feedback loop that perpetuates antibody-driven inflammation and may explain why some patients become refractory to rituximab in SLE.[23]

Following 2NDNR to rituximab, switching to humanised anti-CD20 mAbs restores depletion and response in SLE. Ocrelizumab and ofatumumab are both type 1 anti-CD20 mAbs. The primary endpoint was met in ocrelizumab-treated groups in RA trials[24] and was investigated in SLE.[25] However, development in these indications was halted after an increase in opportunistic infections, some of which fatal were reported.[26] All three patients in our study had major clinical responses and prolonged remission for over 5-year period post-ocrelizumab. Ofatumumab is licenced for resistant chronic lymphocytic leukaemia and has demonstrated efficacy in RA.[27] Both patients in our study responded well to ofatumumab included one who achieved complete depletion for the first time from B-cell depleting therapy. Additionally, a few case series have recently reported on its efficacy in extrarenal and refractory lupus nephritis.[28 29] Alternatively, other anti-CD20 agents with enhanced antibody-dependent cellular cytotoxicity may be more effective in SLE. *In vitro* obinutuzumab demonstrated enhanced depletion was achieved with this type 2 mAb, compared with rituximab.[30]

This study has several limitations. First, an interobserver variability could have occurred in BILAG assessments due to the lengthy follow-up duration and a cohort that was highly heterogeneous in lupus manifestations. However, the BILAG scores reflected the clinician's intention-to-treat, and the patients were managed in a dedicated single centre, thus allowing for consistency in assessment. Second, B-cells and laboratory data were missing in some cases. As these were deemed missing at random, multiple imputation was used to reduce potential bias in parameter estimation as well as enhancing generalisability of the results. Next, concomitant therapy with immunosuppressant were used in more than 60% of the patients, thus efficacy could not be attributed to rituximab alone. Lastly, the lack of control group limits interpretation of efficacy and safety of rituximab.

In conclusion, treatment with anti-CD20 agents can be guided by B-cell monitoring with the aim of achieving complete depletion. About one in eight patients with SLE lose depletion on repeat cycles of rituximab regardless of prior response and secondary non-depletion is associated with anti-rituximab antibodies. Concomitant oral immunosuppressant may help to prevent this. If 2NDNR occurs, switching to humanised anti-CD20 mAbs restores depletion and response. Therefore, alternative anti-CD20 antibodies may be more consistently effective in SLE treatment and several ongoing trials are addressing these issues.

Correction notice This article has been corrected since it published Online First. The abstract has been corrected.

Acknowledgements The authors would like to thank Elizabeth M Hensor for statistical advice, clinicians, pharmacist, study coordinator and lab technicians at the Leeds Connective Tissue Disease and Vasculitis Clinic particularly Mike Martin, Jacqueline Andrews, Maya Buch, Colin Pease, Shouvik Dass, Sinisa Savic, John Bamford, Paul Beirne, Mark Goodfield, Tina Hawkins, Huma Cassamoali, Diane Corscadden and Katie Mbara for their substantial contribution in the acquisition of the data.

Contributors MYMY, PE and EMV: substantial contributions to the conception or design of the work, or the acquisition, analysis or interpretation of data, drafting the work or revising it critically for important intellectual content, final approval of the version published and 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. DS, YME-S, ED and ACR: substantial contributions to the conception or design of the work, or the acquisition, analysis or interpretation of data, drafting the work or revising it critically for important intellectual content and final approval of the version published.

Funding This research was funded/supported by the National Institute for Health Research (NIHR) and NIHR Leeds Biomedical Research Centre based at Leeds Teaching Hospitals NHS Trust; and NIHR Research Grants (DRF-2014-07-155) and (CS-2013-13-032). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests EMV is an NIHR Clinician Scientist. He has received honoraria and research grant support from Roche, GSK and AstraZeneca. PE has received consultant fees from BMS, Abbott, Pfizer, MSD, Novartis, Roche and UCB. He

has received research grants paid to his employer from AstraZeneca, Abbott, BMS, Pfizer, MSD and Roche.

Patient consent The study does not contain any personal medical information about an identifiable living individual, thus patient consent is not required.

Ethics approval The use of rituximab, ofatumumab and ocrelizumab were all approved by Leeds Teaching Hospitals NHS Trust Drug and Therapeutic Committee. Analysis of samples for antirituximabantibody was approved by the Leeds (East) Research Ethics Committee (REC), 10/H1306/88, and the committee confirmed that other aspects of the study did not require ethical approval in accordance with the UK National Health Service REC guidelines.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Aguiar R, Araújo C, Martins-Coelho G, *et al*. Use of Rituximab in systemic lupus erythematosus: a single Center experience over 14 years. *Arthritis Care Res* 2017;69:257-262.
2 Vital EM, Dass S, Buch MH, *et al*. B cell biomarkers of rituximab responses in systemic lupus erythematosus. *Arthritis Rheum* 2011;63:3038–47.
3 Galarza-Maldonado C, Kourilovitch MR, Molineros JE, *et al*. The administration of low doses of rituximab followed by hydroxychloroquine, prednisone and low doses of mycophenolate mofetil is an effective therapy in latin american patients with active systemic lupus erythematosus. *Autoimmun Rev* 2010;10:108–11.
4 Díaz-Lagares C, Croca S, Sangle S, *et al*. Efficacy of rituximab in 164 patients with biopsy-proven lupus nephritis: pooled data from european cohorts. *Autoimmun Rev* 2012;11:357–64.
5 Iaccarino L, Bartoloni E, Carli L, *et al*. Efficacy and safety of off-label use of rituximab in refractory lupus: data from the italian Multicentre registry. *Clin Exp Rheumatol* 2015;33:449–56.
6 Witt M, Grunke M, Proft F, *et al*. Clinical outcomes and safety of rituximab treatment for patients with systemic lupus erythematosus (SLE) - results from a nationwide cohort in Germany (GRAID). *Lupus* 2013;22:1142–9.
7 Ramos-Casals M, Soto MJ, Cuadrado MJ, *et al*. Rituximab in systemic lupus erythematosus: a systematic review of off-label use in 188 cases. *Lupus* 2009;18:767–76.
8 Merrill JT, Neuwelt CM, Wallace DJ, *et al*. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. *Arthritis Rheum* 2010;62:222–33.
9 Rovin BH, Furie R, Latinis K, *et al*. Efficacy and safety of rituximab in patients with active proliferative lupus nephritis: the Lupus Nephritis Assessment with Rituximab study. *Arthritis Rheum* 2012;64:1215–26.
10 Md Yusof MY, Vital EM, Emery P. B-cell-targeted therapies in systemic lupus erythematosus and ANCA-associated vasculitis: current progress. *Expert Rev Clin Immunol* 2013;9:761–72.
11 Reddy V, Croca S, Gerona D, *et al*. Serum rituximab levels and efficiency of B cell depletion: differences between patients with rheumatoid arthritis and systemic lupus erythematosus. *Rheumatology* 2013;52:951–2.
12 Reddy V, Cambridge G, Isenberg DA, *et al*. Internalization of rituximab and the efficiency of B cell depletion in rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Rheumatol* 2015;67:2046–55.
13 Alexander T, Sarfert R, Klotsche J, *et al*. The proteasome inhibitior bortezomib depletes plasma cells and ameliorates clinical manifestations of refractory systemic lupus erythematosus. *Ann Rheum Dis* 2015;74:1474–8.
14 Liu Z, Zou Y, Davidson A. Plasma cells in systemic lupus erythematosus: the long and short of it all. *Eur J Immunol* 2011;41:588–91.
15 Dass S, Rawstron AC, Vital EM, *et al*. Highly sensitive B cell analysis predicts response to rituximab therapy in rheumatoid arthritis. *Arthritis Rheum* 2008;58:2993–9.
16 Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
17 Isenberg DA, Rahman A, Allen E, *et al*. BILAG 2004. Development and initial validation of an updated version of the british Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology* 2005;44:902–6.
18 Yee CS, Cresswell L, Farewell V, *et al*. Numerical scoring for the BILAG-2004 index. *Rheumatology* 2010;49:1665–9.
19 Vital EM, Wittmann M, Edward S, *et al*. Brief report: responses to rituximab suggest B cell-independent inflammation in cutaneous systemic lupus erythematosus. *Arthritis Rheumatol* 2015;67:1586–91.
20 Vital EM, Dass S, Rawstron AC, *et al*. Management of nonresponse to rituximab in rheumatoid arthritis: predictors and outcome of re-treatment. *Arthritis Rheum* 2010;62:1273–9.
21 Vital EM, Dass S, Buch MH, *et al*. An extra dose of rituximab improves clinical response in rheumatoid arthritis patients with initial incomplete B cell depletion: a randomised controlled trial. *Ann Rheum Dis* 2015;74:1195–201.
22 Carter LM, Isenberg DA, Ehrenstein MR. Elevated serum BAFF levels are associated with rising anti-double-stranded DNA antibody levels and disease flare following B cell depletion therapy in systemic lupus erythematosus. *Arthritis Rheum* 2013;65:2672–9.
23 Ehrenstein MR, Wing C. The BAFFling effects of rituximab in lupus: danger ahead? *Nat Rev Rheumatol* 2016;12:367–72.
24 Rigby W, Tony HP, Oelke K, *et al*. Safety and efficacy of ocrelizumab in patients with rheumatoid arthritis and an inadequate response to methotrexate: results of a forty-eight-week randomized, double-blind, placebo-controlled, parallel-group phase III trial. *Arthritis Rheum* 2012;64:350–9.
25 Mysler EF, Spindler AJ, Guzman R, *et al*. Efficacy and safety of ocrelizumab in active proliferative lupus nephritis: results from a randomized, double-blind, phase III study. *Arthritis Rheum* 2013;65:2368–79.
26 Emery P, Rigby W, Tak PP, *et al*. Safety with ocrelizumab in rheumatoid arthritis: results from the ocrelizumab phase III program. *PLoS One* 2014;9:e87379.
27 Østergaard M, Baslund B, Rigby W, *et al*. Ofatumumab, a human anti-CD20 monoclonal antibody, for treatment of rheumatoid arthritis with an inadequate response to one or more disease-modifying antirheumatic drugs: results of a randomized, double-blind, placebo-controlled, phase I/II study. *Arthritis Rheum* 2010;62:2227–38.
28 Thornton CC, Ambrose N, Ioannou Y. Ofatumumab: a novel treatment for severe systemic lupus erythematosus. *Rheumatology* 2015;54:559–60.
29 Haarhaus ML, Svenungsson E, Gunnarsson I. Ofatumumab treatment in lupus nephritis patients. *Clin Kidney J* 2016;9:552–5.
30 Reddy V, Klein C, Isenberg DA, *et al*. Obinutuzumab outperforms rituximab at inducing b-cell cytotoxicity in vitro through fc-mediated effector mechanisms in rheumatoid arthritis and systemic lupus erythematosus. *Arthritis & Rheumatology* 2015;67:1442–5.

EXTENDED REPORT

Repeated administration of dapirolizumab pegol in a randomised phase I study is well tolerated and accompanied by improvements in several composite measures of systemic lupus erythematosus disease activity and changes in whole blood transcriptomic profiles

Chris Chamberlain,[1] Peter J Colman,[1] Ann M Ranger,[2] Linda C Burkly,[3] Geoffrey I Johnston,[1] Christian Otoul,[4] Christian Stach,[5] Miren Zamacona,[1] Thomas Dörner,[6] Murray Urowitz,[7] Falk Hiepe[6]

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-2017-211388).

[1]UCB Pharma, Slough, UK
[2]Unum Therapeutics, Cambridge, Massachusetts, USA
[3]Biogen, Cambridge, Massachusetts, USA
[4]UCB Pharma, Braine-L'Alleud, Belgium
[5]UCB Pharma, Monheim, Germany
[6]Department of Rheumatology and Clinical Immunology, Charité Universitätsmedizin Berlin, Berlin, Germany
[7]University of Toronto, Toronto, Canada

Correspondence to
Professor Chris Chamberlain, UCB Pharma, 208 Bath Road, Slough SL1 3WE, UK; Chris.Chamberlain@ucb.com

Received 1 November 2017
Revised 1 June 2017
Accepted 10 June 2017
Published Online First 11 August 2017

To cite: Chamberlain C, Colman PJ, Ranger AM, *et al*. *Ann Rheum Dis* 2017;**76**:1837–1844.

ABSTRACT

Objectives Systemic lupus erythematosus (SLE) is a heterogeneous autoimmune disease associated with diffuse immune cell dysfunction. CD40–CD40 ligand (CD40L) interaction activates B cells, antigen-presenting cells and platelets. CD40L blockade might provide an innovative treatment for systemic autoimmune disorders. We investigated the safety and clinical activity of dapirolizumab pegol, a polyethylene glycol conjugated anti-CD40L Fab' fragment, in patients with SLE.

Methods This 32-week randomised, double-blind, multicentre study (NCT01764594) evaluated repeated intravenous administration of dapirolizumab pegol in patients with SLE who were positive for/had history of antidouble stranded DNA/antinuclear antibodies and were on stable doses of immunomodulatory therapies (if applicable). Sixteen patients were randomised to 30 mg/kg dapirolizumab pegol followed by 15 mg/kg every 2 weeks for 10 weeks; eight patients received a matched placebo regimen. Randomisation was stratified by evidence of antiphospholipid antibodies. Patients were followed for 18 weeks after the final dose.

Results No serious treatment-emergent adverse events, thromboembolic events or deaths occurred. Adverse events were mild or moderate, transient and resolved without intervention. One patient withdrew due to infection. Efficacy assessments were conducted only in patients with high disease activity at baseline. Five of 11 (46%) dapirolizumab pegol-treated patients achieved British Isles Lupus Assessment Group-based Composite Lupus Assessment response (vs 1/7; 14% placebo) and 5/12 (42%) evaluable for SLE Responder Index-4 responded by week 12 (vs 1/7; 14% placebo). Mechanism-related gene expression changes were observed in blood RNA samples.

Conclusions Dapirolizumab pegol could be an effective biological treatment for SLE. Further studies are required to address efficacy and safety.

Trial registration number NCT01764594.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a multisystem and complex autoimmune disease that results in morbidity, an increased mortality rate and a poor quality of life.[1–3] The pathogenesis involves several mechanisms, including proinflammatory presentation of potential autoantigens by the innate immune system, hyperactivation of the adaptive immune system, formation of pathogenic autoantibodies and the deposition of immune complexes capable of affecting diverse organ systems.[4 5]

CD40 ligand (CD40L, also known as CD154) has been shown to be an important immune-inflammatory modulator and as such it is a credible candidate for pharmacological intervention.[6 7] It is widely expressed on naïve and activated CD4+ T cells and platelets; the receptor, CD40, is expressed constitutively on a wide range of cells including antigen presenting cells and B cells.[8–10] CD40–CD40L interaction has been shown to be essential for normal T cell/B cell functional interactions, including the T cell-dependent humoral immune response, T cell activation of antigen presenting cells, augmentation of CD8+ T cell responses, immunoglobulin (Ig) class-switching and induction of dendritic cell maturation.[10–12] Blockade of CD40–CD40L may decrease the immune activation seen in autoimmune disorders, and this approach has been shown to be efficacious in diverse models of experimental autoimmunity.[13]

In previous studies, the monoclonal anti-CD40L IgG1 antibody, hu5c8 (ruplizumab; BG9588), showed evidence of potential efficacy in patients with lupus nephritis and idiopathic thrombocytopenic purpura;[14 15] however, clinical trials were halted because of a higher than expected occurrence of thromboembolic events.[15 16] It is proposed that the observed treatment-related (TR) thromboembolic events occurred as a result of platelet activation and aggregation, due to the formation of anti-CD40L antibody and soluble CD40L immune complexes that tether to platelets via binding of sCD40L to surface-expressed CD40 and activate platelets through interactions of the Fc with Fc gamma receptor IIA on the platelet surface.[17 18] Dapirolizumab pegol (CDP7657), an anti-CD40L Fab' antibody fragment conjugated to polyethylene

glycol (PEG),[19] was designed to address the potential safety problems caused by the Fc moiety, while retaining favourable pharmacokinetics (PK). In vitro assays have demonstrated that dapirolizumab pegol is a potent antagonist of CD40L binding to CD40, and dose-dependent inhibition of antibody responses with dapirolizumab pegol have been demonstrated in both humanised severe combined immune deficient mice and cynomolgus macaques.[19 20] No histopathological evidence of increased thrombovasculopathy or in vitro platelet activation with dapirolizumab pegol compared with placebo was observed in these models.

We report a phase I study designed to evaluate the safety, tolerability and PK of repeat intravenous dosing of dapirolizumab pegol in patients with active SLE. The study also aimed to assess the effect of dapirolizumab pegol on disease activity and CD40L pathway modulation.

METHODS

Trial design and interventions

A randomised, double-blind, multicentre, placebo-controlled, exploratory phase I study was performed at multiple sites across Europe (Belgium, Bulgaria, Germany, Poland, Romania, Russia and Spain). Patients were randomised (2:1) to receive intravenous-administered dapirolizumab pegol 30 mg/kg followed by dapirolizumab pegol 15 mg/kg every 2 weeks for 10 weeks (total of 5 doses) or to receive a matching placebo regimen (see online supplementary figure S1). Participants were randomised according to evidence of antiphospholipid (aPL) antibodies; these could include historic reports of all known antiphospholipid antibodies. The expected duration of study participation was 32 weeks, comprising an initial 4-week screening period, a 10-week dosing period and an 18-week follow-up period (see online supplementary figure S1). Dosing in the study was intended to achieve a probable therapeutic concentration of dapirolizumab pegol. Data from human studies completed with another anti-CD40L antibody (ruplizumab; BG9588) indicated quantifiable effects on antidouble stranded (ds) DNA antibody levels in patients with SLE[15] when a concentration of >100 µg/mL was achieved (UCB data on file). This was further supported by preclinical data showing a dose-dependent inhibition of antibody response with dapirolizumab pegol in cynomolgus macaques.[19]

Participants

Patients were aged 18–75 years with a diagnosis of SLE satisfying the American College of Rheumatology classification and a Safety of Estrogen in Lupus Erythematosus National Assessment Modification to the Systemic Lupus Erythematosus Disease Activity Index - 2000 (SELENA SLEDAI) score ≥4 at screening.[21 22] All patients had positive anti-dsDNA antibodies (defined as >10 IU/L using an enzyme-labelled anti-isotope assay (ELiA) (Thermo Fisher Scientific: Phadia)) or were antinuclear antibody (ANA) positive (>1:80 using an ELiA assay (Thermo Fisher Scientific: Phadia)) at screening or had a previously documented positive anti-dsDNA antibody or ANA assay.

Patients taking corticosteroids, antimalarial drugs or immunosuppressants were required to be on a stable dose (no greater than 20 mg/day for oral prednisolone or equivalent) for at least 28 days before the first dose of dapirolizumab pegol and were required to maintain this dose for at least 28 days following the start of dapirolizumab pegol administration. There were no restrictions on prior biological therapy use to treat SLE.

Patients with active, neuropsychiatric SLE, active severe glomerulonephritis, an existing or documented acute renal flare in the previous 6 months and/or decreased renal function (estimated glomerular filtration rate <30 mL/min/1.73 m^2; urinary protein >2 g/24 hours or urinary protein:creatinine ratio of >200 mg/mmol) were excluded. Also excluded were patients with a significantly increased risk of thromboembolic events (history of thromboembolism within 1 year prior to screening, vascular graft, valvular heart disease, atrial fibrillation, presence of at least one Factor V Leiden mutant allele, a positive test for aPL antibodies with no stable anticoagulation treatment for at least 28 days prior to the first dose of dapirolizumab pegol).

Clinical endpoints

The primary objective of this study was to evaluate the safety and tolerability of repeated dapirolizumab pegol administration in patients with active SLE. The secondary objective was to assess the PK of dapirolizumab pegol; exploratory objectives monitored disease activity and/or CD40L pathway modulation (via messenger RNA (mRNA) changes and laboratory measures including platelet counts, autoantibodies and complement components) following repeated intravenous administration of dapirolizumab pegol.

Disease activity measures included the British Isles Lupus Assessment Group (BILAG) 2004 index assessment using centralised grading, which reflects disease activity within the last 4 weeks[22 23] and SELENA SLEDAI composite scoring to reflect disease activity within the last 10 days. Responder indices included the BILAG-based Composite Lupus Assessment (BICLA) and the Systemic Lupus Erythematosus Responder Index-4 (SRI-4).[24 25] Physician global assessment was collected via a visual analogue scale ranging from 1–100, and the assessor was also blind to the Subject Global Assessment.

Transcriptional analysis (mRNA signature profiling)

Peripheral blood samples used for gene expression analyses were collected using PAXgene Blood RNA tubes (PreAnalytix GmbH). RNA was isolated using the Agencourt RNAdvance Blood kit (Beckman). cDNA was prepared with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). For quantitative real-time PCR, custom primers and probe sequences were designed using the NCBI Nucleotide website and Primer Express software and ordered from Applied Biosystems (see online supplementary table S1). A panel of genes expressed by B cells and plasma cells was selected for analysis of CD40 pathway modulation. In addition, gene transcripts associated with SLE disease activity, including type I interferon (IFN)-response genes, and those associated with various cellular subsets, were evaluated. All transcripts were normalised to the geometric mean of four housekeeping genes. For each patient, transcript expression for each gene at each timepoint was expressed relative to day 0 (fold change from baseline).

For the analysis, genes were grouped into functional domains associated with either their cellular expression (such as B cells, T cells and natural killer cells) or based on their coordinated expression pattern and association with SLE disease activity (such as IFN-response genes) as described by Chaussabel *et al*.[26] Internal decision criteria to indicate an effect on gene expression were designed to identify consistent changes, recognising the inherent architecture of the data and controlling the false positive rate (see online supplementary material). The prespecified significance level was chosen to achieve an overall false positive rate of 5% when comparing two treatment groups of size 8

and 16 and were finalised prior to assay of any transcriptomic samples.

Randomisation and blinding

A randomisation list was generated by the contract research organisation (CRO) using the Interactive Voice Response/Interactive Web Response Systems. Participants were randomised according to the presence of aPL antibodies. Patients received a five-digit subject number at screening, which was used along with the generated randomisation code to allocate phial numbers to the subject at each treatment. Throughout the study, investigators, study site and CRO staff, with the exception of pharmacy monitors, remained blinded to treatment allocations, as did patients and sponsor staff.

The study was conducted in accordance with International Conference on Harmonisation Good Clinical Practice requirements. The study was run under the aegis of a Data Monitoring Committee formed to monitor the ongoing safety of the study; the study could have been stopped for any significant safety concern, especially the observation of more than one thromboembolic event.

Statistical methods

The full analysis set (FAS) consisted of all randomised patients who received at least one dose of study treatment (dapirolizumab pegol or placebo). The pharmacodynamic per protocol analysis set was a subset of the FAS including those patients who had no important protocol deviations affecting pharmacodynamic variables. The evaluable population for the PK analysis consisted of those patients in the FAS who had no important protocol deviations affecting PK variables and had data from at least one PK sample available.

All statistical analyses were considered exploratory. A statistical analysis plan was completed in advance of database lock and study unblinding. BILAG-BICLA and SRI-4, were implemented to define responders and were analysed separately (see online supplementary material). No escape treatment rule was implemented; no patients received escape treatment. The efficacy analysis was performed in a subset of patients with a baseline SELENA SLEDAI score ≥6 and at least two BILAG grade Bs. The number of responders in each treatment group was calculated for both BICLA and SRI-4, and the responder rates were compared between treatments using Fisher's exact test.

Statistical analysis was performed for measurements of anti-dsDNA, complement C3, complement C4 and for IgG, IgA and the ratio of IgA:IgG. Each variable was subject to a repeated-measures analysis of covariancewith visit as the repeated measure and visit, treatment and visit-by-treatment interaction as fixed effects. Baseline was included as covariate. In the mRNA expression analysis, each transcript type was subject to analysis of variance with subject as random effect and visit, treatment and visit-by-treatment interaction as fixed effects. All original and derived parameters were listed and described using summary statistics. The outputs from these analyses informed the internal decision criteria described above.

RESULTS

Participants

In total 68 patients were screened, 44 were screen failures (majority due to inclusion and exclusion criteria linked to laboratory parameters) and 24 patients were randomised to treatment (dapirolizumab pegol n=16, placebo n=8; figure 1). Baseline characteristics are displayed in table 1 and were broadly similar for each treatment group.

Safety

Based on the primary safety variables, multiple doses of dapirolizumab pegol were well tolerated in patients with SLE. The

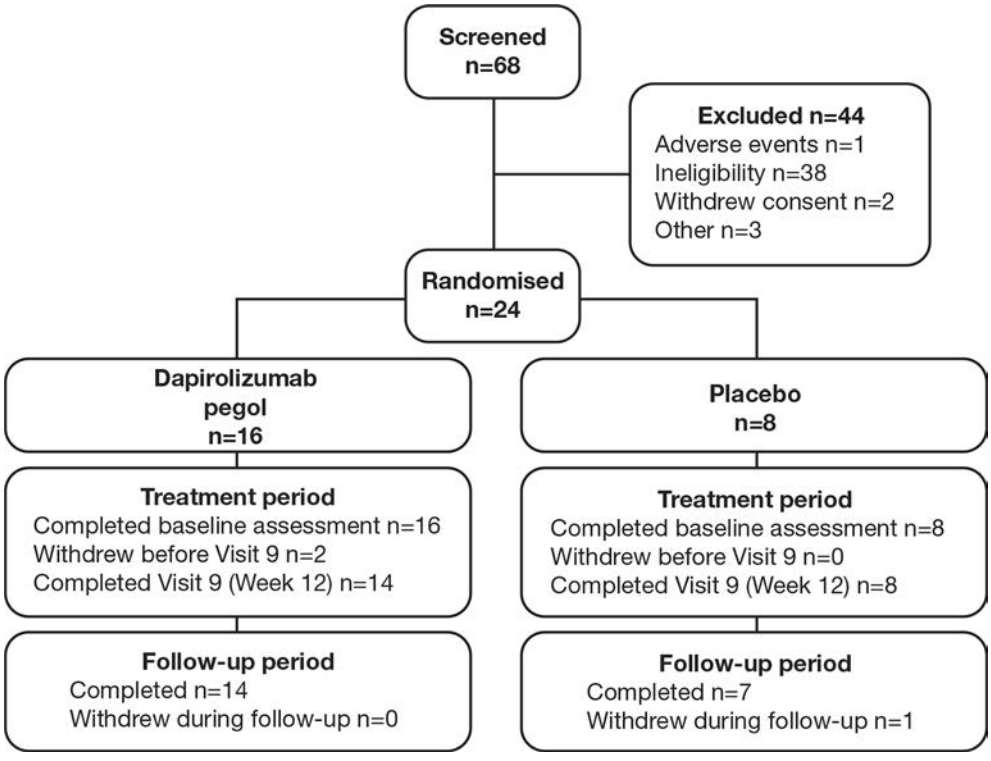


Figure 1 CONSORT diagram.

Table 1 Baseline patient demographics and characteristics

		Placebo	Dapirolizumab pegol
Randomised		**n=8**	**n=16**
Age (years)	Median (range)	38.95 (18.1–59.8)	38.40 (29.5–61.1)
Female	n (%)	8 (100)	13 (81.3)
Weight (kg)	Median (range)	66.5 (42.0–88.0)	60.2 (49.0–105.0)
Height (cm)	Median (range)	162.5 (153.0–172.0)	164.5 (153.0–179.0)
BMI (kg/m^2)	Median (range)	25.2 (17.9–33.2)	23.11 (16.7–42.6)
Racial group – white	n (%)	8 (100)	15 (93.8)
Duration of disease (years)	Median (range)	7.1 (1.0–31.6)	8.9 (0.6–24.7)
SELENA SLEDAI total score	Median (range)	8 (4–14)	9 (2–14)
Physician's Global Assessment of Disease	Median (range)	31.0 (0–68)	28.5 (0–67)
Subject's Global Assessment of Disease	Median (range)	36.0 (0–72)	32.5 (0–75)
GFR (mL/min/1.73 m^2)	Median (range)	75.0 (57–111)	80.0 (60–157)
At least 1 BILAG grade A	n (%)	1 (12.5)	4 (25.0)
At least 1 BILAG grade B	n (%)	7 (87.5)	12 (75.0)
BILAG total score	Median (range)	10.0 (2–21)	13.0 (2–21)
Anti-dsDNA antibody (IU/mL)*	Median (range)	16.0 (1–90)	10.0 (1–475)
Anti-dsDNA antibody†	Positive n (%)	4 (50)	7 (43.8)
	Negative n (%)	3 (37.5)	8 (50)
Antiphospholipid antibodies	Present (n)	1	7
	Absent (n)	7	9
Complement C3 (mg/L)	Median (range)	1065 (380–1360)	840 (460–1460)
Complement C4 (mg/L)‡	Median (range)	215 (40–330)	145 (30–420)
IgG (g/L)	Median (range)	10.8 (4.7–13.8)	12.9 (5.6–19.6)
Prior exposure			
Immunosuppressant	n (%)	3 (37.5)	4 (25.0)
Antimalarials	n (%)	6 (75.0)	3 (18.8)
Corticosteroid #	n (%)	7 (87.5)	15 (93.8)
Concomitant medications			
Immunosuppressants	n (%)	1 (12.5)	2 (12.5)
Antimalarials	n (%)	7 (87.5)	7 (43.8)
Corticosteroids	n (%)	7 (87.5)	14 (87.5)

*Defined as >10 IU/L using an enzyme-labelled anti-isotope assay (ELiA).
†Results >15 IU/mL were considered positive, 10–15 IU/mL indeterminate and <10 IU/mL were considered negative.
‡Normal range 100–400 mg/L.
#Median total daily dose 10.0 mg.
BMI, body mass index; BILAG, British Isles Lupus Assessment Group; GFR, glomerular filtration rate; SELENA SLEDAI, Safety of Estrogen in Lupus Erythematosus National Assessment Modification to the Systemic Lupus Erythematosus Disease Activity Index – 2000.

overall incidence of treatment-emergent adverse events (TEAEs) in the dapirolizumab pegol group was 87.5% (14/16 patients) compared with 62.5% in the placebo group (5/8 patients; see online supplementary table S2). All TEAEs in the dapirolizumab pegol group were considered mild or moderate in intensity; none were severe. No serious AEs were reported, and no patients died during the study. No TEAEs related to thromboembolic events or laboratory findings suggestive of thromboembolic events were reported in either the dapirolizumab pegol or placebo group.

The most commonly reported TEAEs in the dapirolizumab pegol group were nasopharyngitis (6/16; 37.5%), headache (4/16; 25%), upper respiratory tract infection (3/16; 18.8%), anaemia and diarrhoea (2/16; 12.5% for each) (table 2). One subject (12.5%) in the placebo group reported severe TEAEs of musculoskeletal pain and neck pain. One subject (6.3%) in the dapirolizumab pegol group discontinued the study because of an upper respiratory tract infection; the event was deemed mild and not considered by the investigator to be related to treatment. There was a numerically greater incidence in TEAEs related to infection in the dapirolizumab pegol group (11/16; 68.8%) compared with the placebo group (3/8; 37.5%). However, none of the infections were serious or considered opportunistic. There were no noteworthy TEAEs related to infusion reactions.

Four (25.0%) dapirolizumab pegol-treated patients and three (37.5%) placebo-treated patients were considered to have treatment-related (TR)-TEAEs, as determined by the investigator. In the dapirolizumab pegol group, these included nasopharyngitis (2/16; 12.5%), diarrhoea, herpes simplex, paronychia, increased lipase, hypernatraemia, headache and dysuria (1/16; 6.3%, each). In the placebo group, TR-TEAEs included feeling hot, musculoskeletal pain, neck pain and hot flush (1/8; 12.5% each). There were no clinically significant abnormalities and no substantive differences between the dapirolizumab pegol and placebo groups in laboratory values, including platelet, lymphocyte and neutrophil counts, vital signs or ECGs. There were no increases in the doses of concomitant corticosteroids, antimalarial drugs or immunosuppressants during the study.

Disease activity

Overall, there was greater improvement in the clinical measures of disease activity in the dapirolizumab pegol group than in the placebo group. Improvements from baseline were observed for

Table 2 Incidence of TEAEs occurring in ≥2 patients in any category

	Placebo			Dapirolizumab pegol		
MedDRA (V.17.0)	**Male**	**Female**	**Total**	**Male**	**Female**	**Total**
SOC	**n=0**	**n=8**	**n=8**	**n=3**	**n=13**	**n=16**
Preferred Term	**n (%) (#)**	**n (%) (#)**	**n (%) (#)**	**n (%) (#)**	**n (%) (#)**	**n (%) (#)**
Blood and lymphatic system disorders						
Anaemia	0	0	0	0	2 (15.4) (2)	2 (12.5) (2)
Gastrointestinal disorders						
Diarrhoea	0	0	0	0	2 (15.4) (2)	2 (12.5) (2)
Nausea	0	1 (12.5) (3)	1 (12.5) (3)	0	1 (7.7) (1)	1 (6.3) (1)
Infections and infestations						
Nasopharyngitis	0	0	0	1 (33.3) (1)	5 (38.5) (5)	6 (37.5) (6)
Upper respiratory tract infection	0	1 (12.5) (1)	1 (12.5) (1)	1 (33.3) (1)	2 (15.4) (2)	3 (18.8) (3)
Nervous system disorders						
Headache	0	1 (12.5) (6)	1 (12.5) (6)	0	4 (30.8) (5)	4 (25.0) (5)

(#) Number of individual occurrences of the TEAE in that category.
MedDRA, Medical Dictionary for Regulatory Activities; SOC, system organ class; TEAE, treatment-emergent adverse event.

the SELENA SLEDAI (figure 2), Subject's Global Assessment of Disease, and BILAG total score. There was a greater proportion of BICLA (5/11; 45.5%) and SRI-4 (5/12; 41.7%) responders in the dapirolizumab pegol group at week 12, compared with the placebo group (1/7; 14.3% for both BICLA and SRI-4).

Exploratory analyses

At the mRNA transcript level, expression changes in genes within the plasma cell and B cell domains were noted, consistent with modulation of the CD40 pathway. Among the plasma cell genes, the dapirolizumab pegol group exhibited more rapid and greater decreases in the expression of several Ig-associated genes (secretory IgA, IgG, Igκ, Igλ and J chain) starting at week 2 and maintained over the treatment period, compared with placebo (figure 3A). Among the B cell genes, the dapirolizumab pegol group showed a transient increase in CD19 and CD20 RNA transcripts at week 2 (figure 3B). Several of the patients in the dapirolizumab pegol group exhibited a >2-fold reduction in the expression of type I IFN-response genes (MX1 (n=6), OAS1 (n=4), IFITM3 (n=9), G1P2 (n=5)) for at least two time points (figure 4). These changes were more dramatic in the responder group. No other consistent changes in other functional group expression patterns were observed (data not shown).

There was a treatment difference in anti-dsDNA antibody concentrations (p=0.0168) at week 12; however, there were only seven patients in the dapirolizumab pegol group and four

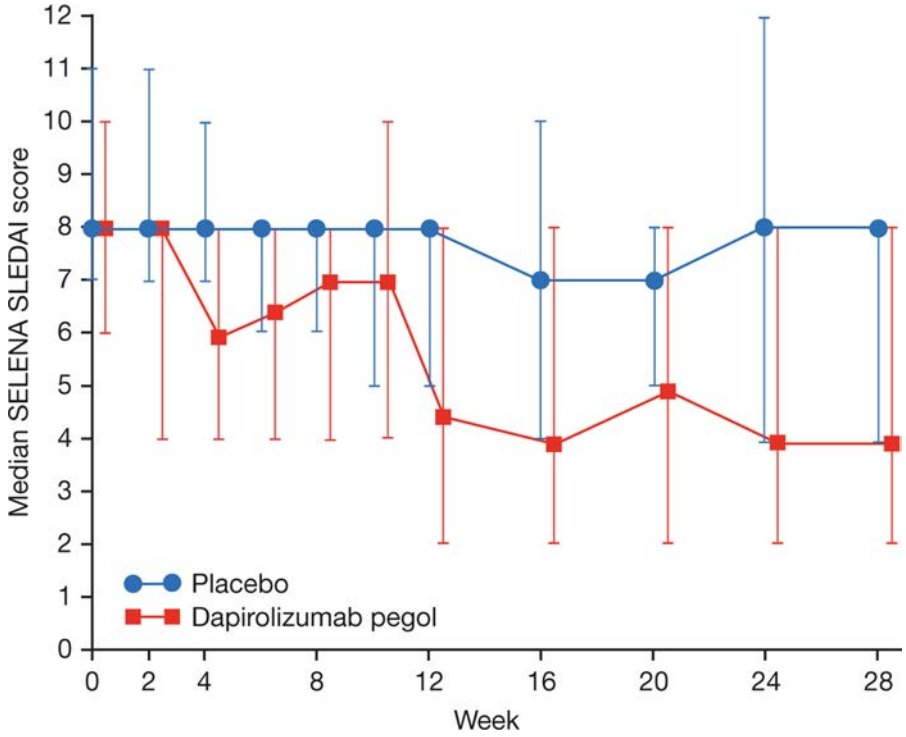


Figure 2 Median SELENA SLEDAI total score over time (PD-PPS). Vertical bars represent the IQRs, that is, the central 50% of the observed data. PD-PPS, pharmacodynamic per protocol analysis set; SELENA SLEDAI, Safety of Estrogen in Lupus Erythematosus National Assessment Modification to the Systemic Lupus Erythematosus Disease Activity Index – 2000.

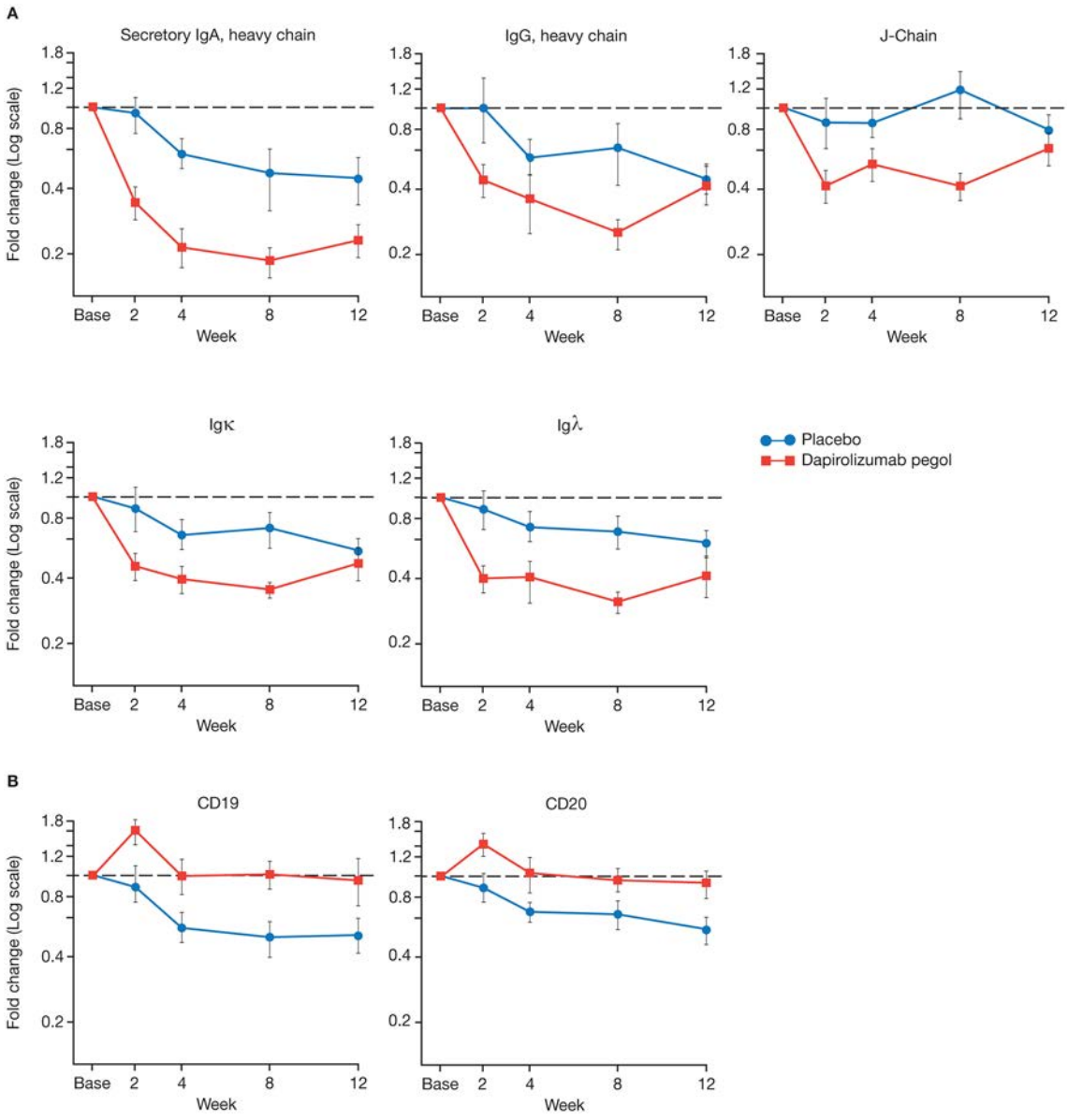


Figure 3 Mean fold change of mRNA levels from baseline for representative genes belonging to the plasma cell domain (A) and B cell domain (B). Plots display arithmetic mean (±SEM). Dapirolizumab pegol-treated (n=16) and placebo-treated (n=8) patients are shown in red and blue, respectively. Dashed line represents no change.

patients in the placebo group with elevated titres (>15 IU/mL) at baseline. No signals were detected in other laboratory measures, including anti-C1q, aPL and complements C3 and C4. In the dapirolizumab pegol group, a reduction in IgG concentrations (p=0.0436) and small reductions in total Ig and IgA concentrations were observed at week 12 compared with the placebo group at the same time point.

PK analysis

PK assessments showed that the exposure of dapirolizumab pegol was maintained with a trough concentration >100 µg/mL (online supplementary figure 2); the doses selected for the study were based on the trough level at which anti-dsDNA antibody decrease was observed in the hu5c8 study.[15] The estimated half-life of dapirolizumab pegol was in the range of 7.8–14 days, geometric mean AUC_{w0-10} was 22 326 day*µg/mL and geometric mean C_{max} at week 10 was 582.2 µg/mL. The PK of dapirolizumab pegol was not affected by the presence of antidrug or anti-PEG antibodies.

DISCUSSION

Studies in animal models have shown that blockade of CD40L is efficacious in inflammatory and autoimmune conditions.[19 20] While caution is needed when extrapolating such data to humans, CD40L blockade could be an innovative approach for the treatment of SLE.[27] Dapirolizumab pegol is a purified recombinant, humanised Fab' antibody fragment covalently bound to PEG that targets CD40L. The primary objective of this study was to evaluate the safety and tolerability of repeated intravenous doses of dapirolizumab pegol in patients with mild to moderate SLE. The secondary objective was to assess PK; exploratory objectives included evaluation of the effects of dapirolizumab pegol on disease activity, biomarkers of disease activity and CD40L pathway modulation.

In previous clinical studies of CD40L inhibition using monoclonal full-length IgG1 antibodies, the incidence of thromboembolic events has been higher than expected.[15 16 28] The Fc portion of the full-length antibody has a critical role in the mechanism leading to thromboembolic events.[17 18] Since dapirolizumab

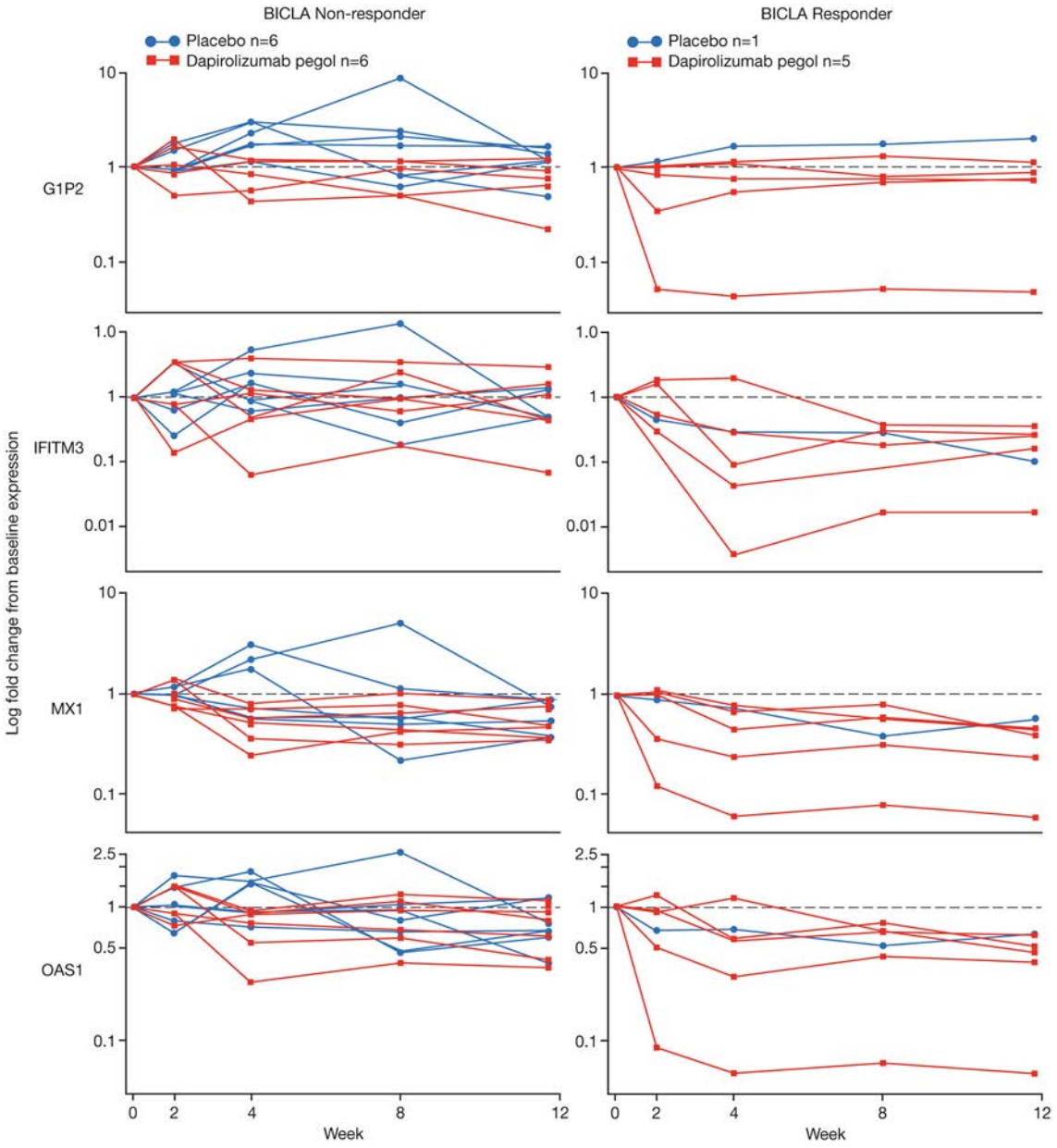


Figure 4 Mean fold change in the RNA transcript levels of four type I interferon-response genes for BICLA responders and non-responders in dapirolizumab pegol-treated (red) and placebo-treated (blue) patients. Dashed line represents no change from baseline level.

pegol lacks the Fc portion, it would not be expected to be associated with increased thromboembolic risk, a hypothesis supported by the absence of thromboembolic events in the present study and a previous non-human primate study in Rhesus monkeys.[19] Multiple doses of dapirolizumab pegol were well tolerated. There was a higher incidence of non-serious infection in the dapirolizumab pegol group. No TEAEs related to thromboembolic events or laboratory findings suggestive of thromboembolic events were reported during the study, and there was no evidence of enhanced procoagulatory effects. A trend towards an increased infection rate was noted in the dapirolizumab pegol arm, but this was not accompanied by changes in total lymphocyte or neutrophil counts. These safety findings are comparable with those from a previous single-dose, double-blind, first-in-human, phase I study of dapirolizumab pegol (NCT01093911).[29] None of the study subjects experienced any dose limiting toxicities and no thromboembolic events were reported.[29]

In terms of clinical measures of disease activity, there was greater improvement in the dapirolizumab pegol group compared with the placebo group. Improvements from baseline were observed in SELENA SLEDAI, Subject's Global Assessment of Disease, BILAG total score, BICLA and SRI-4 responders. In addition, statistically significant changes were observed in expression of genes associated with B cell and plasma cell function, as were reductions in the expression of IFN-response genes, consistent with known functions of CD40L.

The promising safety and preliminary efficacy findings reported here must be interpreted with caution as this small, exploratory study was not powered to demonstrate statistical significance on all outcomes reported. Only one dose level was evaluated, so no dose effects were studied; however, a phase II dose-ranging study to better define the optimal therapeutic dose and regimen is underway (NCT02804763).

This study has shown that dapirolizumab pegol, administered in multiple doses over 12 weeks, appears well tolerated in patients with mild to moderate SLE with no major safety concerns. We observed gene transcription changes associated with inhibition of the CD40–CD40L interaction and improvement in clinical measures of disease activity in the dapirolizumab pegol group compared with the placebo group. These results support further investigation of dapirolizumab pegol as a novel treatment for SLE.

Acknowledgements The study was funded by UCB Pharma and Biogen. The authors would like to acknowledge Norm Allaire, Julie Czerkowicz and other Biogen colleagues for their contributions to the gene expression analysis; Huo Li and Carrie Wager of Biogen for normalisation of the qPCR data. The authors would like to acknowledge Laura Griffin, PhD, of iMed Comms, an Ashfield Company, part of UDG Healthcare plc for medical writing support that was funded by UCB Pharma and Biogen. The authors would like to thank the investigators, patients and their families who participated in this study.

Contributors CC, AMR, LCB and MZ contributed to the conception and design of the study; CC, TD, FH, GIJ, CO and MU were involved in the acquisition of data; CC, AMR, PJC, TD, FH, GIJ, CO, CS, MU, LCB and MZ contributed to the analysis and interpretation of data. All authors contributed to drafting and/or revising the manuscript.

Funding UCB Pharma and Biogen.

Competing interests CC, PJC, GIJ, CO and MZ are full-time employees of UCB Pharma and hold stock awards and/or options. CS is a full-time employee of UCB Biosciences GmbH and holds stock awards and/or options. AMR was a full-time employee and stock holder of Biogen at the time the study was conducted. LCB is a full time employee and stockholder of Biogen. FH has received consultancy fees from UCB, Sanofi, Eli Lilly, Baxter, BMS and research grants from Deutsche Forschungsgemeinschaft, IMI (PRECISESADS) and attended speakers' bureau for GSK, Roche Pharma and Pfizer. TD has received consultancy fees and research grants from UCB, Biogen, Roche, Sanofi, Eli Lilly, Jansen and research grants from Deutsche Forschungsgemeinschaft and EU Horizon 2020 (Harmonics). MU has served as Chair of the Data and Safety Monitoring Committee for the study.

Patient consent A patient consent form was completed.

Ethics approval This was a national, regional, Independent Ethics Committee or Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Bernatsky S, Boivin JF, Joseph L, *et al*. Mortality in systemic lupus erythematosus. *Arthritis Rheum* 2006;54:2550–7.
2 Arnaud L, Fagot JP, Mathian A, *et al*. Prevalence and incidence of systemic lupus erythematosus in France: a 2010 nation-wide population-based study. *Autoimmun Rev* 2014;13:1082–9.
3 Devilliers H, Amoura Z, Besancenot JF, *et al*. LupusQoL-FR is valid to assess quality of life in patients with systemic lupus erythematosus. *Rheumatology* 2012;51:1906–15.
4 Gottschalk TA, Tsantikos E, Hibbs ML. Pathogenic inflammation and its therapeutic targeting in systemic lupus erythematosus. *Front Immunol* 2015;6:550.
5 Yaniv G, Twig G, Shor DB, *et al*. A volcanic explosion of autoantibodies in systemic lupus erythematosus: a diversity of 180 different antibodies found in SLE patients. *Autoimmun Rev* 2015;14:75–9.
6 Bankert KC, Oxley KL, Smith SM, *et al*. Induction of an altered CD40 signaling complex by an antagonistic human monoclonal antibody to CD40. *J Immunol* 2015;194:4319–27.
7 Sidiropoulos PI, Boumpas DT. Lessons learned from anti-CD40L treatment in systemic lupus erythematosus patients. *Lupus* 2004;13:391–7.
8 Elgueta R, Benson MJ, de Vries VC, *et al*. Molecular mechanism and function of CD40/CD40L engagement in the immune system. *Immunol Rev* 2009;229:152–72.
9 Lesley R, Kelly LM, Xu Y, *et al*. Naive CD4 T cells constitutively express CD40L and augment autoreactive B cell survival. *Proc Natl Acad Sci U S A* 2006;103:10717–22.
10 Aloui C, Prigent A, Sut C, *et al*. The signaling role of CD40 ligand in platelet biology and in platelet component transfusion. *Int J Mol Sci* 2014;15:22342–64.
11 Schönbeck U, Libby P. The CD40/CD154 receptor/ligand dyad. *Cell Mol Life Sci* 2001;58:4–43.
12 Ström L, Laurencikiené J, Miskinienė A, *et al*. Characterization of CD40-dependent immunoglobulin class switching. *Scand J Immunol* 1999;49:523–32.
13 Toubi E, Shoenfeld Y. The role of CD40-CD154 interactions in autoimmunity and the benefit of disrupting this pathway. *Autoimmunity* 2004;37:457–64.
14 Kalunian KC, Davis JC, Merrill JT, *et al*. Treatment of systemic lupus erythematosus by inhibition of T cell costimulation with anti-CD154: a randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2002;46:3251–8.
15 Boumpas DT, Furie R, Manzi S, *et al*. A short course of BG9588 (anti-CD40 ligand antibody) improves serologic activity and decreases hematuria in patients with proliferative lupus glomerulonephritis. *Arthritis Rheum* 2003;48:719–27.
16 Kawai T, Andrews D, Colvin RB, *et al*. Thromboembolic complications after treatment with monoclonal antibody against CD40 ligand. *Nat Med* 2000;6:114.
17 Langer F, Ingersoll SB, Amirkhosravi A, *et al*. The role of CD40 in CD40L- and antibody-mediated platelet activation. *Thromb Haemost* 2005;93:1137–46.
18 Robles-Carrillo L, Meyer T, Hatfield M, *et al*. Anti-CD40L immune complexes potently activate platelets in vitro and cause thrombosis in FCGR2A transgenic mice. *J Immunol* 2010;185:1577–83.
19 Shock A, Burkly L, Wakefield I, *et al*. CDP7657, an anti-CD40L antibody lacking an fc domain, inhibits CD40L-dependent immune responses without thrombotic complications: an in vivo study. *Arthritis Res Ther* 2015;17:234.
20 Wakefield P I, Burkl C, *et al*. CDP7657, a Monovalent Fab PEG Anti-CD40L antibody, inhibits immune responses in both HuSCID mice and Non-Human Primates. [abstract]. *Arthritis Rheum* 2010;62(Suppl 10):1245.
21 Petri M, Buyon J, Skovron M, *et al*. Disease activity and health status (SF-36) in post-menopausal systemic lupus erythematosus: the SELENA trial. *Arthritis Rheum* 1997;40:1062.
22 Romero-Diaz J, Isenberg D, Ramsey-Goldman R. Measures of adult systemic lupus erythematosus: updated version of british Isles Lupus Assessment Group (BILAG 2004), European Consensus Lupus Activity Measurements (ECLAM), Systemic lupus activity measure, revised (SLAM-R), Systemic lupus activity questionnaire for Population studies (SLAQ), Systemic lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K), and systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index (SDI). *Arthritis Care Res* 2011;63(Suppl 11):S37–46.
23 Isenberg DA, Rahman A, Allen E, *et al*. BILAG 2004. Development and initial validation of an updated version of the british Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology* 2005;44:902–6.
24 Wallace DJ. Evaluation of treatment success in systemic lupus Erythematosus clinical trials: development of the british Isles Lupus Assessment Group-based Composite Lupus Assessment Endpoint. *Arthritis Rheum* 2011;61:1143–51.
25 Furie RA, Petri MA, Wallace DJ, *et al*. Novel evidence-based systemic lupus erythematosus responder index. *Arthritis Rheum* 2009;61:1143–51.
26 Chaussabel D, Quinn C, Shen J, *et al*. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity* 2008;29:150–64.
27 van Kooten C, Banchereau J. CD40-CD40 ligand. *J Leukoc Biol* 2000;67:2–17.
28 Liossis SN, Sfikakis PP. Costimulation blockade in the treatment of rheumatic diseases. *BioDrugs* 2004;18:95–102.
29 Tocoian A, Buchan P, Kirby H, *et al*. First-in-human trial of the safety, pharmacokinetics and immunogenicity of a PEGylated anti-CD40L antibody fragment (CDP7657) in healthy individuals and patients with systemic lupus erythematosus. *Lupus* 2015;24:1045–56.

OPEN ACCESS

EXTENDED REPORT

Oral contraceptives, breastfeeding and the risk of developing rheumatoid arthritis: results from the Swedish EIRA study

Cecilia Orellana,[1] Saedis Saevarsdottir,[1,2] Lars Klareskog,[2] Elizabeth W Karlson,[3] Lars Alfredsson,[1,4] Camilla Bengtsson[1]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211620).

[1]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[2]Rheumatology Unit, Department of Medicine, Solna, Karolinska University Hospital and Karolinska Institutet, Stockholm, Sweden
[3]Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[4]Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden

Correspondence to
Cecilia Orellana, Institute of Environmental Medicine, Karolinska Institutet, Box 210, S1171 77 Stockholm, Sweden; cecilia.orellana@ki.se

Received 10 April 2017
Revised 15 June 2017
Accepted 24 June 2017
Published Online First 28 July 2017

ABSTRACT

Objectives To study whether oral contraceptive (OC) use or breastfeeding (BF) influence the risk of rheumatoid arthritis (RA), stratifying the cases by presence/absence of anticitrullinated protein antibodies (ACPA), and whether these factors interact with known risk factors in the development of ACPA-positive RA.

Methods Women aged ≥18 years, participants in the population-based case–control Swedish Epidemiological Investigation of RA study (2641 cases/4251 controls), completed an extensive questionnaire regarding OC, BF and potential confounders. We calculated ORs, with 95% CIs, adjusted for age, residential area, smoking and alcohol consumption. Attributable proportion due to interaction (AP) was estimated to evaluate presence of interaction.

Results Compared with never users, ever and past OC users had a decreased risk of ACPA-positive RA (OR=0.84 (95% CI 0.74 to 0.96); OR=0.83 (95% CI 0.73 to 0.95), respectively). No significant associations were found for ACPA-negative RA. Long duration of OC use (>7 years vs never use) decreased the risk of both ACPA-positive (p=0.0037) and ACPA-negative RA (p=0.0356). A history of long BF decreased the risk only of ACPA-positive RA in a dose-dependent manner (p=0.0086), but this trend did not remain after adjustments. A significant interaction was observed between the lack of OC use and smoking (AP=0.28 (95% CI 0.14–0.42)) on the risk of ACPA-positive RA. No interactions were found for BF.

Conclusions OC decreased the risk of RA, especially ACPA-positive RA, where an interaction with smoking was observed. A long duration of OC use decreased the risk of both disease subsets. We could not confirm an association between BF and a decreased risk of either ACPA-positive or ACPA-negative RA.

INTRODUCTION

Rheumatoid arthritis (RA) is among the most common autoimmune diseases, with a complex interplay of genetic and environmental factors involved in its aetiology.[1 2] Since the disease is two to three times more common among women as compared with men,[3–5] it has been suggested that hormonal and reproductive factors might partly explain this sex difference.

Regarding oral contraceptive (OC) use and the risk of RA, some studies have shown an inverse association,[6–11] but the majority of reports have been unable to demonstrate an association.[12–23] Only a few previous reports have taken seropositivity into account, either exploring the classic rheumatoid factor (RF)[6 9 11 15 21] or presence/absence of anticitrullinated protein antibodies (ACPA).[12 14] Furthermore, disparate results so far might be explained by methodological issues, such as the use of prevalent cases for analysis,[14] inclusion of non-population controls[9 11] or relatively few cases.[11 12 21]

Breastfeeding (BF) has been associated with a decreased risk of RA,[13 15 24 25] and a long duration of BF seems to have the strongest association.[15 24] However, some studies have found an increased RA risk.[12 26] Analyses taking seropositivity into account have yielded disparate results.[12 13 15 26] Among these, Berglin *et al* reported that a longer BF history provided a higher risk of RA among those carrying the *PTPN22* 1858T variant or were positive for ACPA or RF.[12] Apart from these studies, the influence of BF on ACPA-positive/ACPA-negative RA has not been further investigated.

For the ACPA-positive subgroup of RA, several risk factors have been identified, including smoking, the *PTPN22*R620W* (1858 C/T) risk allele and the *HLA-DRB1* shared epitope (SE) allele.[1 27–31] In contrast, for the ACPA-negative subgroup of RA, only a few risk factors have been identified.[2 31] ACPA-status and the classic RF highly correlate, and risk factors for seropositive/negative RA behave similarly.[2 30 32]

The aim of this study was to investigate the association between both OC use and total history of BF among parous women, and the risk of developing RA stratifying the cases by ACPA-status (positive/negative), using data from a large population-based case–control study. Moreover, the aim was to explore potential additive interactions between BF and OC, respectively, in regard to known risk factors for ACPA-positive disease, namely smoking status, presence of SE alleles and *PTPN22* gene.

METHOD

Study design

This study was based on data from the Swedish Epidemiological Investigation of RA (EIRA) comprising women above 18 years, living in defined geographical areas of Sweden, between 1996 and 2014. The general design of the EIRA study has been described in detail elsewhere.[33] Incident cases of RA were diagnosed by rheumatologists and included if they fulfilled either the American College of Rheumatology 1987 criteria[34] or the

To cite: Orellana C, Saevarsdottir S, Klareskog L, *et al*. *Ann Rheum Dis* 2017;**76**:1845–1852.

latest 2010 RA criteria.[35] Twenty-four cases were diagnosed according to the new criteria alone. Controls were randomly selected from the national population register and matched to the cases by age (5-year group) and residential area. For further details, see online supplementary text (online supplementary file 1). All participants provided written informed consent, and ethical approval was obtained from the Regional Ethical Review Board at Karolinska Institutet, Stockholm, Sweden.

Data collection

Participants completed an extensive questionnaire regarding lifestyle and environmental exposures, including OC use, BF and potential confounders. Information about OC use was available for the entire study period, whereas information on BF history among parous women was only available from 2006.

Between 1996 and 2014, a total of 2809 cases and 5312 controls were identified; of these, 2676 cases (95%) and 4251 controls (80%) answered the questionnaire. Blood samples were available from all participating cases.

Antibody assays and genotyping

Blood samples were assayed for ACPA-status using the Immunoscan-RA Mark2 ELISA test (Euro-Diagnostica, Malmö, Sweden).[36 37] The cut-off value for ACPA-positive RA was 25 U/mL. A total of 35 and 13 cases lacking information on ACPA-status were excluded from the OC and BF analyses, respectively.

Genotyping of the *PTPN22* and *HLA-DRB1* genes was conducted as previously described.[38 39] Among *HLA-DRB1* genes, *DRB1*01*, *DRB1*04* and *DRB1*10* genes were defined as SE alleles. Any genotype containing 1 or 2 of these genes was considered as having 'any SE allele', versus those not having any of the genes ('no SE alleles').

Exposures

The year in which the first symptoms of RA occurred was defined as the index-year for each case. Controls were then assigned the same index-year as their matched case.

Current users of OCs were defined as those who were currently using OCs during the index-year and who had started at least the year before index-year. Participants who started OC use during index-year (four cases/seven controls) and those with missing information on OC use (59 cases/115 controls) were excluded from the analyses. Past users were defined as those who used OCs in the past and had stopped at least the year before the index-year. Ever users were defined as current and past users while never users were women who had not used OCs at any time before the index-year.

Parous women were defined as those who had given birth before or during the index-year. Total BF history among parous women was calculated as the sum of the duration of BF for each child born and categorised as 0–6, 7–12 and ≥13 months, according to quartile distribution among controls. Participants with missing information on BF history (78 cases/148 controls) were excluded from analyses. Parous women who did not breast feed (two cases/14 controls) were included in the reference category.

Statistical analysis

Odds ratios (OR) with 95% confidence intervals (CI) of RA overall, ACPA-positive and ACPA-negative RA, associated with OC use and BF were calculated by means of unconditional logistic regression. Regarding OC use, current/past/ever users were compared with never users. Duration of OC use was categorised according to the median value among controls (≤7/>7 years). For the BF analyses, the shortest duration of BF (0–6 months) was used as the reference category.

All analyses were adjusted for the matching variables (age and residential area). We conducted additional adjustments (each variable was investigated separately) for parity (yes/no), number of children (1, 2, 3 and ≥4), body mass index (<25/≥25 kg/m^2), menopausal status, use of postmenopausal hormone therapy (ever/never), age at menarche (≤11, 12, 13 and ≥14 years), age at first birth (<22, 22–24, 25–29 and >29 years), time between last delivered child and the index-year (0–24, 25–30, 31–37 and >37 years), index-year intervals, university education (yes/no), pack-years of cigarette smoking (0–<10, ≥10–<20 and ≥20) and alcohol consumption (low (including non-drinkers), medium and high). We also adjusted for OC use when analysing BF as the main exposure and vice versa. Only smoking and alcohol consumption made a change in the ORs and were retained in the final analyses.

Potential interaction was estimated using departure from additivity of effects (additive interaction), as suggested by Rothman.[40] We tested for interactions in the same manner for both OC use and BF with well-established risk factors of RA: smoking, SE alleles and *PTPN22* risk allele.

To evaluate interaction, the attributable proportion due to interaction (AP) was calculated together with the 95% CI.[41] The AP is the proportion of the incidence among people exposed to two interacting factors, indicating their joint effect apart from the sum of their independent effects. For further details, see online supplementary text.

All analyses were carried out using the Statistical Analysis System (SAS) V.9.4.

RESULTS

In total, 2641 cases and 4251 controls were available for the OC analyses. Overall, 1756 (66.5%) cases were ACPA-positive and the mean time between symptom onset and diagnosis was 10 months for both ACPA subsets. A total of 2578 cases and 4129 controls were included in the OC analyses after all exclusions. For BF, a total of 1242 cases and 2658 controls were available for analysis (for the period 2006–2014), of which 884 cases and 1949 controls where parous women with available BF history. Baseline characteristics of participants are presented in table 1.

OC use and risk of RA

Ever users of OCs had a decreased odds of developing RA overall compared with never users (OR=0.87, 95% CI 0.78 to 0.97). The OR for current and past users were 0.85 (95% CI 0.68 to 1.06) and 0.87 (95% CI 0.78 to 0.98), respectively. The association between ever and past OC use was significant for ACPA-positive, but not for the smaller subset of ACPA-negative RA, and remained significant after adjustment for pack-years of smoking and alcohol consumption (table 2).

A longer duration of ever OC use (above the median value of 7 years) was associated with a decreased risk of RA overall (OR=0.81, 95% CI 0.71 to 0.92). The trend with a longer duration was significant for both ACPA-positive (p=0.0037) and ACPA-negative RA (p=0.0356). Similar result was observed for past OC use except for ACPA-negative RA, probably due to lack of power (table 3). Separate analyses for OC using RF yielded similar results (data not shown).

Table 1 Characteristics of cases and controls. EIRA, Sweden, 1996–2014

	Cases (n=2641) N (%), mean±SD		Controls (n=4251) N (%), mean±SD
	ACPA-positive RA 1756 (66.5%)	ACPA-negative RA 885 (33.5%)	
Age at inclusion (years)	50.9±13.0	52.0±13.5	51.4±13.4
Age at menarche (years)	13.2±1.4	13.2±1.4	13.1±1.5
Parous	1375 (78.3)	718 (81.1)	3376 (79.4)
Number of children	2.2±1.2	2.2±0.8	2.2±0.9
Age at first birth (years)	24.8±4.9	24.5±4.9	25.6±5.0*
Age at menopause (years)	49.6±5.6	49.8±5.3	50.0±5.4
Oral contraceptive use†			
Ever	1135 (64.7)	582 (65.8)	2862 (67.4)‡
Current	134 (7.6)	61 (6.9)	331 (7.8)
Past	1001 (57.1)	521 (58.9)	2531 (59.6)‡
Never	572 (32.6)	289 (32.7)	1267 (29.9)
Missing	46 (2.6)	13 (1.5)	115 (2.7)
Breast feeding (months)§			
None	1 (0.1)	1 (0.3)	14 (0.7)
1–6	193 (28.7)	80 (27.6)	519 (24.7)
7–12	192 (28.6)	83 (28.6)	574 (27.4)
≥13	234 (34.8)	100 (34.5)	842 (40.1)
Missing	52 (7.7)	26 (9.0)	148 (7.1)
Total duration of breast feeding (months) according to parity§			
One child	6.4±5.6	4.8±2.7	6.9±5.4
Two children	11.7±8.2	11.7±8.1	12.4±8.3
Three children or more	22.5±18.1	19.5±12.9	20.7±13.9
Ever use of PMH¶	117 (26.3)	67 (29.7)	412 (29.5)
BMI ≥25 kg/m^2	749 (42.7)	409 (46.2)	1704 (40.1)‡
University degree	469 (26.7)	251 (28.4)	1425 (33.5)*
Ever smoker	1175 (66.9)	531 (60.0)	2266 (53.3)*
Pack-years			
Never smokers	571 (32.5)	348 (39.3)	1943 (45.7)
0–10	367 (20.9)	185 (21.0)	963 (22.7)‡
10–20	316 (18.0)	132 (14.9)	531 (12.5)*
≥20	409 (23.3)	149 (16.8)	530 (12.5)*
Other	82 (4.7)	63 (7.1)	243 (5.7)‡
Missing	11 (0.6)	8 (0.9)	41 (0.9)
Alcohol consumption			
Non-drinkers	213 (12.1)	100 (11.3)	330 (7.8)*
Low	892 (50.9)	418 (47.3)	1991 (46.9)
Moderate	408 (23.3)	228 (25.8)	1045 (24.6)
High	235 (13.4)	138 (15.6)	864 (20.4)*
Missing	5 (0.3)	0 (0)	14 (0.3)

Baseline characteristics among participants who replied to the questionnaire, excluding cases lacking ACPA-status (35 cases).
Information on age at menarche and age at menopause available for 1211 cases/2596 controls and 757 cases/1548 controls, respectively.
A pack-year is defined as 20 cigarettes smoked every day for 1 year. The category 'Other' includes those smoking other tobacco than cigarettes (eg, cigarillos, cigars or pipe tobacco).
Alcohol consumption defined as number of drinks per week (one drink=12 g of alcohol) and categorised according to the quartile distribution among the controls. The two lowest categories (non-drinkers and low consumption) were merged for analyses.
*p Value <0.0001 for the difference between cases and controls.
†Oral contraceptive use after exclusion of four cases/seven controls who initiated use during the index-year. Ever oral contraceptive use is the sum of current and past use.
‡p Value <0.05 for the difference between cases and controls.
§Information on breastfeeding available for 884 cases and 1949 controls (all parous women) from 2006. Quartile distribution among controls, with the two highest categories merged into one.
¶Only among postmenopausal women.
ACPA, anticitrullinated protein antibodies; BMI, body mass index; EIRA, Epidemiological Investigation of Rheumatoid Arthritis; PMH, postmenopausal hormone therapy; RA, rheumatoid arthritis.

BF and risk of RA

Compared with women who breast fed for 0–6 months, those who breast fed their children for 7–12 months had an OR of 0.93 (95% CI 0.75 to 1.14) of developing RA overall, whereas BF for 13 months or more had an OR of 0.77 (95% CI 0.63 to 0.94). This declining trend was statistically significant for

Table 2 ORs of developing RA overall and ACPA-positive/ACPA-negative RA according to oral contraceptive use. EIRA, Sweden, 1996–2014

ACPA status	Oral contraceptive use*	Ca/Co	OR (95% CI)†	OR (95% CI)‡
RA overall	Ever	1717/2862	0.87 (0.78 to 0.97)	0.87 (0.78 to 0.98)
	Current	195/331	0.85 (0.68 to 1.06)	0.89 (0.71 to 1.12)
	Past	1522/2531	0.87 (0.78 to 0.98)	0.87 (0.78 to 0.98)
	Never	861/1267	1.0	1.0
	Missing	59/115	–	–
ACPA-positive	Ever	1135/2862	0.84 (0.74 to 0.95)	0.84 (0.74 to 0.96)
	Current	134/331	0.86 (0.67 to 1.11)	0.92 (0.71 to 1.19)
	Past	1001/2531	0.84 (0.74 to 0.95)	0.83 (0.73 to 0.95)
	Never	572/1267	1.0	1.0
	Missing	46/115	–	–
ACPA-negative	Ever	582/2862	0.94 (0.80 to 1.10)	0.93 (0.79 to 1.10)
	Current	61/331	0.83 (0.59 to 1.17)	0.81 (0.57 to 1.16)
	Past	521/2531	0.95 (0.80 to 1.12)	0.94 (0.79 to 1.11)
	Never	289/1267	1.0	1.0
	Missing	13/115	–	–

*Participants who started OC use during index-year (four cases/seven controls) were excluded from the analysis. Ever is the sum of current and past OC users.
†Adjusted for age and residential area.
‡Adjusted for age, residential area, smoking (pack-years) and alcohol consumption (low (including non-drinkers), medium and high).
ACPA, anticitrullinated protein antibodies; Ca/Co, number of cases/controls; EIRA, Epidemiological Investigation of Rheumatoid Arthritis; OC, oral contraceptive; RA, rheumatoid arthritis.

ACPA-positive, but not for ACPA-negative RA. These estimates were attenuated after adjustment for smoking and alcohol consumption (table 4). Analyses using RF instead of ACPA gave similar results (data not shown).

Interaction analyses

Never OC use among never smokers was not associated with risk of ACPA-positive RA (OR=0.99, 95% CI 0.81 to 1.21). Compared with never smoking women which had used OCs, women who had smoked and used OCs had an OR=1.71 (95% CI 1.47 to 1.99), whereas women who had smoked and never used OCs had an OR=2.34 (95% CI 1.95 to 2.82) (table 5). Moreover, a significant interaction on the additive scale was found between smoking and never use of OCs (AP=0.28, 95% CI 0.14 to 0.42) regarding the risk of ACPA-positive RA, indicating that among smokers the risk was more pronounced in never OC users than in ever OC users. No significant interactions were found between OC use and SE alleles, the *PTPN22* gene or between BF and any of the three factors explored (data not shown).

DISCUSSION

In this large population-based case–control study of incident RA, with careful matching between cases and controls and extensive exposure information, we found that women who had ever used OCs had a significantly decreased risk of developing RA. The estimates were similar for current and past use, although only significant in the larger group of past users. When stratifying by ACPA-status, the association was only significant for ACPA-positive RA in both crude and adjusted models. A significant

Table 3 ORs of developing RA overall and ACPA-positive/ACPA-negative RA according to duration of oral contraceptive use. EIRA, Sweden, 1996–2014

		Ever OC use		Current OC use		Past OC use	
ACPA-status	Duration of OC use*	Ca/Co	OR (95% CI)†	Ca/Co	OR (95% CI)†	Ca/Co	OR (95% CI)†
RA overall	Never	852/1245	1.0	852/1245	1.0	852/1245	1.0
	≤7 years	865/1348	0.94 (0.83 to 1.07)	59/85	1.16 (0.77 to 1.76)	806/1263	0.93 (0.82 to 1.06)
	>7 years	835/1481	0.81 (0.71 to 0.92)	134/242	0.99 (0.74 to 1.33)	701/1239	0.81 (0.71 to 0.93)
	p-trend		0.0014		0.9982		0.0021
ACPA-positive	Never	565/1245	1.0	565/1245	1.0	565/1245	1.0
	≤7 years	556/1348	0.89 (0.76 to 1.03)	39/85	1.18 (0.73 to 1.90)	517/1263	0.88 (0.75 to 1.02)
	>7 years	570/1481	0.80 (0.69 to 0.93)	95/242	0.95 (0.69 to 1.32)	475/1239	0.80 (0.68 to 0.93)
	p-trend		0.0037		0.8011		0.0039
ACPA-negative	Never	287/1245	1.0	287/1245	1.0	287/1245	1.0
	≤7 years	309/1348	1.04 (0.86 to 1.25)	20/85	1.15 (0.61 to 2.18)	289/1263	1.04 (0.86 to 1.25)
	>7 years	265/1481	0.82 (0.67 to 0.99)	39/242	1.09 (0.68 to 1.74)	226/1239	0.83 (0.67 to 1.01)
	p-trend		0.0356		0.7056		0.0636

26 cases and 55 controls lacked information on duration of oral contraceptive use.
*Duration of OC use categorised according to median value among controls.
†Adjusted for age, residential area, smoking (pack-years) and alcohol consumption (low (including non-drinkers), medium and high).
ACPA, anticitrullinated protein antibodies; Ca/Co, number of cases/controls; EIRA, Epidemiological Investigation of Rheumatoid Arthritis; OC, oral contraceptives; RA, rheumatoid arthritis.

Table 4 ORs of developing RA overall and ACPA-positive/ACPA-negative RA according to breastfeeding. EIRA, Sweden, 2006–2014

ACPA-status	Breastfeeding*	Ca/Co	OR (95% CI)†	OR (95% CI)‡
RA overall	≤6 months	275/533	1.0	1.0
	7–12 months	275/574	0.93 (0.75 to 1.14)	0.99 (0.80 to 1.23)
	≥13 months	334/842	0.77 (0.63 to 0.94)	0.88 (0.71 to 1.08)
	Missing	78/148	–	–
	p-value trend	–	0.0075	0.1919
ACPA-positive	≤6 months	194/533	1.0	1.0
	7–12 months	192/574	0.91 (0.72 to 1.15)	0.99 (0.78 to 1.26)
	≥13 months	234/842	0.74 (0.59 to 0.93)	0.86 (0.68 to 1.09)
	Missing	52/148	–	–
	p-value trend	–	0.0086	0.2096
ACPA-negative	≤6 months	81/533	1.0	1.0
	7–12 months	83/574	0.97 (0.70 to 1.35)	1.01 (0.72 to 1.42)
	≥13 months	100/842	0.83 (0.60 to 1.15)	0.91 (0.65 to 1.27)
	Missing	26/148	–	–
	p-value trend	–	0.2405	0.5446

*Breastfeeding duration categorised according to quartiles values among controls, merging the two highest categories.
†Adjusted for age and residential area.
‡Adjusted for age, residential area, smoking (pack-years) and alcohol consumption (low (including non-drinkers), medium and high).
ACPA, anticitrullinated protein antibodies; Ca/Co, number of cases/controls; EIRA, Epidemiological Investigation of Rheumatoid Arthritis; RA, rheumatoid arthritis.

dose–response association was observed for duration of ever OC use both for ACPA-positive and ACPA-negative RA. Non-use of OC significantly interacted with smoking regarding the risk of ACPA-positive RA. Furthermore, BF also decreased the risk of RA in a dose-dependent manner (total duration), but this trend was only significant for ACPA-positive RA and did not maintain after adjustments.

EIRA has the advantage of being one of the largest population-based case–control studies comprising incident cases of RA with concordant information on environmental and genetic factors. The selection of controls (randomly and continuously from the same study base as the cases) minimises the possible selection bias in this step. Another major strength of our study was the possibility to adjust our results with respect to several potential confounders.

Several limitations of our study should be mentioned. First, although the participation proportion among controls was high (80%), selection bias may have occurred if the controls did not

Table 5 ORs of developing ACPA-positive RA for subjects exposed to OC and ever smoking/HLA-DRB1 SE alleles/PTPN22 in women aged 18 years or above. EIRA, Sweden, 1996–2014

OC use*	Smoking	Ca/Co	OR (95% CI)†	OR (95% CI)‡
Ever	Never	358/1205	1.0	1.0
Never	Never	201/684	1.05 (0.86 to 1.28)	0.99 (0.81 to 1.21)
Ever	Ever	771/1632	1.61 (1.39 to 1.87)	1.71 (1.47 to 1.99)
Never	Ever	364/563	2.33 (1.94 to 2.80)	2.34 (1.95 to 2.82)
AP§	–	–	0.29 (0.15 to 0.43)	0.28 (0.14 to 0.42)
OC use*	**SE alleles**	**Ca/Co**	**OR (95% CI)†**	**OR (95% CI)‡**
Ever	None	113/449	1.0	1.0
Never	None	60/215	1.24 (0.86 to 1.77)	1.26 (0.87 to 1.83)
Ever	Any	657/531	4.99 (3.93 to 6.33)	5.11 (4.00 to 6.54)
Never	Any	348/243	6.62 (5.03 to 8.70)	6.28 (4.73 to 8.34)
AP§			0.21 (0.04 to 0.38)	0.14 (–0.05 to 0.34)
OC use*	**PTPN22 alleles**	**Ca/Co**	**OR (95% CI)†**	**OR (95% CI)‡**
Ever	None	578/840	1.0	1.0
Never	None	311/389	1.32 (1.09 to 1.60)	1.27 (1.04 to 1.55)
Ever	Any	249/239	1.50 (1.22 to 1.85)	1.53 (1.23 to 1.90)
Never	Any	115/108	1.84 (1.37 to 2.47)	1.76 (1.30 to 2.39)
AP§			0.0007 (–0.32 to 0.33)	–0.02 (–0.37 to 0.33)

*Since ever OC use was associated with a decreased risk of ACPA-positive RA, the risk category included non-OC users for each interaction analysis, which was separately conducted for smoking, SE alleles and *PTPN22*.
†Adjusted for matching variables (age and residential area) and alcohol consumption.
‡Adjusted for matching variables (age and residential area), pack-years of smoking and alcohol consumption (low (including non-drinkers), medium and high).
§The AP estimates the proportion of the excess risk that is due to the interaction *per se* (factor A + factor B) according to the formula $RR_{AB} - RR_A - RR_B + 1 / R_{RAB}$ (where RR=relative risk).
ACPA, anticitrullinated protein antibodies; AP, attributable proportion due to interaction; Ca/Co, number of cases/controls; EIRA, Epidemiological Investigation of Rheumatoid Arthritis; OC, oral contraceptives; RA, rheumatoid arthritis; SE, shared epitope.

reflect the exposure frequency in the study base. However, both BF and ever OC use among controls were very similar to the high frequency of BF[42] and reported OC use[43] among Swedish women, respectively. Second, we did not have detailed information regarding OC preparations or doses, being only able to conduct analyses on OC use as a whole.

Regarding OC use, our finding of a decreased risk of developing RA is in accordance with previous reports.[6–11] Although most previous studies have not observed a statistically significant association,[12–21] some results have suggested a protective effect, but the sample size might have been inadequate to reach definite conclusions. Only borderline associations have been observed in a few recent meta-analyses.[22 23] Our results are in agreement with those from Doran *et al*,[6] who reported a decreased risk among ever (OR=0.57 (95% CI 0.35 to 0.91) but not among current (OR=1.0 (95% CI 0.4 to 2.52) OC users. Another case–control study performed in Sweden showed a non-significant association for ever (OR=0.70, 95% CI 0.40 to 1.24) and current (OR=1.21, 95% CI 0.58 to 2.52) OC users, but the association for past users was significant (OR=0.37, 95% CI 0.16 to 0.86).[8] These findings are in line with our results, although they used old criteria (year 1958) for RA diagnosis. In line with our findings, Berglin *et al* found a protective effect with OC use >7 years.[12]

Previous reports taking seropositivity into account have yielded contradictory results.[6 9 11 12 14 15 21] Doran *et al* found a protective effect of OC exposure on the risk of RF-positive (OR=0.36, 95% CI 0.18 to 0.72) but not RF-negative RA (OR=0.982, 95% CI 0.46 to 2.10).[6] By contrast, Pedersen *et al* reported an increased risk of ACPA-positive RA among ever users of OC (OR=1.65, 95% CI 1.06 to 2.57).[14] However, the inclusion of prevalent cases (diagnosed within 5 years) might entail bias. Our result was notably mainly restricted to ACPA-positive RA and estimates only slightly modified after adjustments. Similar results using RF instead of ACPA-status corroborate the high correlation between ACPA-status and the classic RF.

The current knowledge on the association between BF and RA has not reached firm conclusions. In a large cohort study, Karlson *et al* found a decreased risk of RA among women who breast fed for more than 12 months, with a significant trend with increased duration of BF.[15] Restricting the analyses to RF positive patients, a similar reduction was found for a total BF duration of ≥24 months. In line with these results, a study conducted in Sweden showed a decreased risk of RA among women who breast fed their children for more than a year, with similar trends for RF-negative/RF-positive RA.[13] A similar result was obtained by a recent cross-sectional study in an Asian population.[24] Our estimates were attenuated after adjustments for smoking and alcohol consumption, which indicates their role as important confounders in this study. Analyses using RF yielded similar results as those for ACPA-status. However, in a nested case–control study, Berglin *et al* found a strong association between BF and later development of RA (OR=4.8, 95% CI 1.43 to 15.8) with a higher risk with increasing time of BF and greater among ACPA-positive cases.[12] Similar findings were reported by Brennan and Silman, with a higher risk for RF-positive RA among women who breast fed.[26] These opposite results might be explained by methodological issues (small number of cases (70) and recruitment via a media campaign, respectively). Finally, a recently published systematic review and meta-analysis reported a decreased risk of RA, whether with a longer or shorter duration of BF.[25] Our study confirms and extends these findings by adding the stratification according to ACPA-status, which to our knowledge has not been explored using a large dataset as in our present study.

To the best of our knowledge, no previous study has found evidence of interaction between OC and/or BF and smoking habits or major genetic risk factors of RA, respectively. The significant interaction between lack of OC use and smoking indicates that the risk of ACPA-positive RA associated with smoking is higher among women who never used OCs than among those who did. However, since both smoking and the use of OC have been linked to an increased predisposition to venous thrombotic events (VTE), women with a history of VTEs (especially if they smoke) might be recommended not to use OC by their physician. We can therefore not exclude the possibility that our findings on an interaction between non-OC use and smoking merely reflects that smoking women, who have an increased RA risk, do not receive OC prescription as often. The physiopathology of RA is complex and not fully understood, but our findings may contribute to the knowledge regarding mechanisms of importance for the development of RA.

The postpartum period soon after delivery has been described as a time of higher risk for the onset of RA.[19] The immunostimulating effect of the hormone prolactin, levels of which are elevated during BF, might explain this increased risk immediately after childbirth.[44] However, recent findings indicate that prolactin might act more as a regulator of inflammation, with protective and regenerative functions.[45] Since elevated prolactin levels do not support our findings, a potential biological mechanism might be a prolonged anti-inflammatory effect of progesterone. It has been shown that elevated progesterone levels during pregnancy remain high during BF through expression of progesterone receptors in lymphocytes.[46] Finally, another potential mechanism might be an anti-inflammatory effect of cortisol, which has been found to be significantly higher among postmenopausal women with a history of BF.[47]

The potential effect of hormones contained in OC preparations might vary according to dose and type. Although such information was not available in the present study, the protective effect seems to differ between ACPA-subsets and with a longer duration of OC use, supporting the hypothesis of a dose–response effect.

The protective effect of OC use on the risk of ACPA-positive RA is in line with our previous finding of a reduced risk of ACPA-positive RA, among women who used postmenopausal hormone therapy.[48] On the other hand, the finding of a protective effect of BF on the risk of ACPA-positive (in the crude model), but not ACPA-negative RA, is in line with our previous finding of a risk of ACPA-negative but not ACPA-positive RA during the postpartum period.[49] All of these findings together support the notion of RA as two different disease entities with different risk factors patterns.

In summary, we found an inverse relationship between OC use and the subsequent development of RA, especially ACPA-positive RA. An interaction between never OC use and smoking was also observed for this subgroup of disease, implying that among smokers, the risk was more pronounced in never OC users than in ever OC users. A trend was observed for longer duration of BF and decreased risk of ACPA-positive RA, although not significant after adjustments. In this large population-based study, we were able to address these questions more thoroughly than has been possible before, by examining disease subsets separately, in the context of other risk factors and by considering many potential confounders. Further research is required to explore the biological mechanisms behind our findings and whether hormonal factors have different impact on the ACPA-subsets of RA.

Acknowledgements We would like to thank Marie-Louise Serra and Lena Nise for excellent assistance with the collection of data. We would also like to thank all cases and controls who participated in the study as well as the clinicians and nurses in the EIRA study group.

Contributors CO conducted the statistical analyses and drafted the manuscript; CB initiated the study and was responsible for the analysis, interpreting the results and revising the manuscript; LA, LK and CB contributed to study design; and LA, LK, EWK and SS contributed to data interpretation and critical revision of the manuscript for important intellectual content. All authors have read and approved the final manuscript.

Funding This study was supported by grants from the Swedish Medical Research Council, the Swedish Research Council for Health, Working Life and Welfare, King Gustav Vâ€™s 801-year foundation, Vinnova, the Swedish Foundation for Strategic Research, the Swedish Rheumatic Foundation, the Stockholm County Council, the Insurance Company AFA, the Innovative Medicines Initiative 1 supported BTCure project and the National Institutes of Health (NIH, AR047782). The funding sources had no role in the reporting of the study or in the decision to submit the manuscript for publication.

Competing interests No, there are no competing interests.

Patient consent Obtained.

Ethics approval This study was approved by the Regional Ethical Review Board at Karolinska Institutet, Stockholm, Sweden.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Karlson EW, Deane K. Environmental and gene-environment interactions and risk of rheumatoid arthritis. *Rheum Dis Clin North Am* 2012;38:405–26.
2 Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet* 2009;373:659–72.
3 Humphreys JH, Verstappen SM, Hyrich KL, *et al*. The incidence of rheumatoid arthritis in the UK: comparisons using the 2010 ACR/EULAR classification criteria and the 1987 ACR classification criteria. results from the Norfolk Arthritis Register. *Ann Rheum Dis* 2013;72:1315–20.
4 Kvien TK, Uhlig T, Ødegård S, *et al*. Epidemiological aspects of rheumatoid arthritis: the sex ratio. *Ann N Y Acad Sci* 2006;1069:212–22.
5 Doran MF, Pond GR, Crowson CS, *et al*. Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a forty-year period. *Arthritis Rheum* 2002;46:625–31.
6 Doran MF, Crowson CS, O'Fallon WM, *et al*. The effect of oral contraceptives and estrogen replacement therapy on the risk of rheumatoid arthritis: a population based study. *J Rheumatol* 2004;31:207–13.
7 Reckner Olsson A, Skogh T, Wingren G. Comorbidity and lifestyle, reproductive factors, and environmental exposures associated with rheumatoid arthritis. *Ann Rheum Dis* 2001;60:934–9.
8 Allebeck P, Ahlbom A, Ljungström K, *et al*. Do oral contraceptives reduce the incidence of rheumatoid arthritis? A pilot study using the Stockholm County medical information system. *Scand J Rheumatol* 1984;13:140–6.
9 Vandenbroucke JP, Valkenburg HA, Boersma JW, *et al*. Oral contraceptives and rheumatoid arthritis: further evidence for a preventive effect. *Lancet* 1982;2:839–42.
10 Koepsell T, Dugowson C, Voigt L, *et al*. Preliminary findings from a case-control study of the risk of rheumatoid arthritis in relation to oral contraceptive use. *Br J Rheumatol* 1989;28 Suppl 1:41–5. discussion.
11 Hazes JM, Dijkmans BC, Vandenbroucke JP, *et al*. Reduction of the risk of rheumatoid arthritis among women who take oral contraceptives. *Arthritis Rheum* 1990;33:173–9.
12 Berglin E, Kokkonen H, Einarsdottir E, *et al*. Influence of female hormonal factors, in relation to autoantibodies and genetic markers, on the development of rheumatoid arthritis in northern Sweden: a case-control study. *Scand J Rheumatol* 2010;39:454–60.
13 Pikwer M, Bergström U, Nilsson JA, *et al*. Breast feeding, but not use of oral contraceptives, is associated with a reduced risk of rheumatoid arthritis. *Ann Rheum Dis* 2009;68:526–30.
14 Pedersen M, Jacobsen S, Klarlund M, *et al*. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res Ther* 2006;8:R133.
15 Karlson EW, Mandl LA, Hankinson SE, *et al*. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? results from the nurses' Health Study. *Arthritis Rheum* 2004;50:3458–67.
16 Merlino LA, Cerhan JR, Criswell LA, *et al*. Estrogen and other female reproductive risk factors are not strongly associated with the development of rheumatoid arthritis in elderly women. *Semin Arthritis Rheum* 2003;33:72–82.
17 Pope JE, Bellamy N, Stevens A. The lack of associations between rheumatoid arthritis and both nulliparity and infertility. *Semin Arthritis Rheum* 1999;28:342–50.
18 Brennan P, Bankhead C, Silman A, *et al*. Oral contraceptives and rheumatoid arthritis: results from a primary care-based incident case-control study. *Semin Arthritis Rheum* 1997;26:817–23.
19 Silman A, Kay A, Brennan P. Timing of pregnancy in relation to the onset of rheumatoid arthritis. *Arthritis Rheum* 1992;35:152–5.
20 Spector TD, Roman E, Silman AJ. The pill, parity, and rheumatoid arthritis. *Arthritis Rheum* 1990;33:782–9.
21 del Junco DJ, Annegers JF, Luthra HS, *et al*. Do oral contraceptives prevent rheumatoid arthritis? *JAMA* 1985;254:1938–41.
22 Qi S, Xin R, Guo W, *et al*. Meta-analysis of oral contraceptives and rheumatoid arthritis risk in women. *Ther Clin Risk Manag* 2014;10:915–23.
23 Chen Q, Jin Z, Xiang C, *et al*. Absence of protective effect of oral contraceptive use on the development of rheumatoid arthritis: a meta-analysis of observational studies. *Int J Rheum Dis* 2014;17:725–37.
24 Adab P, Jiang CQ, Rankin E, *et al*. Breastfeeding practice, oral contraceptive use and risk of rheumatoid arthritis among chinese women: the Guangzhou Biobank Cohort Study. *Rheumatology* 2014;53:860–6.
25 Chen H, Wang J, Zhou W, *et al*. Breastfeeding and risk of Rheumatoid Arthritis: a Systematic Review and Metaanalysis. *J Rheumatol* 2015;42:1563–9.
26 Brennan P, Silman A. Breast-feeding and the onset of rheumatoid arthritis. *Arthritis Rheum* 1994;37:808–13.
27 Källberg H, Ding B, Padyukov L, *et al*. Smoking is a Major preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. *Ann Rheum Dis* 2011;70:508–11.
28 Padyukov L, Seielstad M, Ong RT, *et al*. A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis. *Ann Rheum Dis* 2011;70:259–65.
29 Raychaudhuri S. Recent advances in the genetics of rheumatoid arthritis. *Curr Opin Rheumatol* 2010;22:109–18.
30 Klareskog L, Stolt P, Lundberg K, *et al*. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum* 2006;54:38–46.
31 Lahiri M, Morgan C, Symmons DP, *et al*. Modifiable risk factors for RA: prevention, better than cure? *Rheumatology* 2012;51:499–512.
32 Terao C, Ohmura K, Ikari K, *et al*. Effects of smoking and shared epitope on the production of anti-citrullinated peptide antibody in a japanese adult population. *Arthritis Care Res* 2014;66:1818–27.
33 Bengtsson C, Berglund A, Serra ML, *et al*. Non-participation in EIRA: a population-based case-control study of rheumatoid arthritis. *Scand J Rheumatol* 2010;39:344–6.
34 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.
35 Aletaha D, Neogi T, Silman AJ, *et al*. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.
36 Rantapää-Dahlqvist S. Diagnostic and prognostic significance of autoantibodies in early rheumatoid arthritis. *Scand J Rheumatol* 2005;34:83–96.
37 Rönnelid J, Wick MC, Lampa J, *et al*. Longitudinal analysis of citrullinated protein/peptide antibodies (anti-CP) during 5 year follow up in early rheumatoid arthritis: anti-cp status predicts worse disease activity and greater radiological progression. *Ann Rheum Dis* 2005;64:1744–9.
38 Padyukov L, Silva C, Stolt P, *et al*. A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. *Arthritis Rheum* 2004;50:3085–92.
39 Olerup O, Zetterquist H. HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. *Tissue Antigens* 1992;39:225–35.
40 Rothman KJ, ed. *Epidemiology: an introduction*. 2nd ed. USA: Oxford University Press, 2012.
41 Hosmer DW, Lemeshow S. Confidence interval estimation of interaction. *Epidemiology* 1992;3:452–6.
42 Sweden OSo. Breast-feeding and smoking habits among parents of infants born in 2011. 2013. [cited 2015 23 March] http://www.socialstyrelsen.se/publikationer2013/2013-9-18.
43 Josefsson A, Wiréhn AB, Lindberg M, *et al*. Continuation rates of oral hormonal contraceptives in a cohort of first-time users: a population-based registry study, Sweden 2005-2010. *BMJ Open* 2013;3:e003401.
44 Orbach H, Shoenfeld Y. Hyperprolactinemia and autoimmune diseases. *Autoimmun Rev* 2007;6:537–42.

45 Costanza M, Binart N, Steinman L, *et al*. Prolactin: a versatile regulator of inflammation and autoimmune pathology. *Autoimmun Rev* 2015;14:223–30.
46 Szekeres-Bartho J, Barakonyi A, Par G, *et al*. Progesterone as an immunomodulatory molecule. *Int Immunopharmacol* 2001;1:1037–48.
47 Lankarani-Fard A, Kritz-Silverstein D, Barrett-Connor E, *et al*. Cumulative duration of breast-feeding influences cortisol levels in postmenopausal women. *J Womens Health Gend Based Med* 2001;10:681–7.
48 Orellana C, Saevarsdottir S, Klareskog L, *et al*. Postmenopausal hormone therapy and the risk of rheumatoid arthritis: results from the Swedish EIRA population-based case-control study. *Eur J Epidemiol* 2015;30:449–57.
49 Orellana C, Wedrén S, Källberg H, *et al*. EIRA Study Group. Parity and the risk of developing rheumatoid arthritis: results from the Swedish Epidemiological Investigation of Rheumatoid Arthritis study. *Ann Rheum Dis* 2014;73: 752–5.

OPEN ACCESS

EXTENDED REPORT

Patient-reported outcomes from a phase 3 study of baricitinib versus placebo or adalimumab in rheumatoid arthritis: secondary analyses from the RA-BEAM study

Edward C Keystone,[1] Peter C Taylor,[2] Yoshiya Tanaka,[3] Carol Gaich,[4] Amy M DeLozier,[4] Anna Dudek,[5] Jorge Velasco Zamora,[6] Jose Arturo Covarrubias Cobos,[7] Terence Rooney,[4] Stephanie de Bono,[4] Vipin Arora,[4] Bruno Linetzky,[4] Michael E Weinblatt[8]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211259).

[1]The Rebecca MacDonald Centre for Arthritis, Mt. Sinai Hospital, Toronto, Canada
[2]Kennedy Institute of Rheumatology, University of Oxford, Oxford, UK
[3]University of Occupational and Environmental Health, Fukuoka, Japan
[4]Eli Lilly and Company, Indianapolis, Indiana, USA
[5]Medica Pro Familia, Warsaw, Poland
[6]Instituto Médico CER, Buenos Aires, Argentina
[7]Unidad Reumatologica Las Americas S.C.P. Médico, Mexico, USA
[8]Brigham and Women's Hospital, Boston, Massachusetts, USA

Correspondence to
Professor Edward C Keystone, The Rebecca MacDonald Centre for Arthritis and Autoimmunity, Mt Sinai Hospital, University of Toronto, Toronto, Canada; ed.keystone@sinaihealthsystem.ca

Received 3 February 2017
Revised 22 June 2017
Accepted 24 June 2017
Published Online First 10 August 2017

To cite: Keystone EC, Taylor PC, Tanaka Y, *et al*. *Ann Rheum Dis* 2017;**76**:1853–1861.

ABSTRACT

Background To assess the effect of baricitinib on patient-reported outcomes (PROs) in patients with active rheumatoid arthritis and an inadequate response to methotrexate (MTX).

Methods In this double-blind phase 3 study, patients were randomised 3:3:2 to placebo (n=488), baricitinib 4 mg once daily (n=487), or adalimumab 40 mg biweekly (n=330) with background MTX. PROs included the SF-36, EuroQol 5-D (EQ-5D) index scores and visual analogue scale, Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F), Health Assessment Questionnaire-Disability Index (HAQ-DI), Patient's Global Assessment of Disease Activity (PtGA), patient's assessment of pain and Work Productivity and Activity Impairment Questionnaire-Rheumatoid Arthritis (WPAI-RA), and measures collected in electronic patient daily diaries: duration and severity of morning joint stiffness (MJS), Worst Ttiredness and Worst Joint Pain. The primary study endpoint was at week 12. Treatment comparisons were assessed with logistic regression for categorical measures or analysis of covariance for continuous variables.

Results Compared with placebo and adalimumab, baricitinib showed statistically significant improvements ($p \leq 0.05$) in HAQ-DI, PtGA, pain, FACIT-F, SF-36 physical component score, EQ-5D index scores and WPAI-RA daily activity at week 12. Improvements were maintained for measures assessed to week 52. Statistically significant improvement in patient diary measures (MJS duration and severity), worst tiredness and worst joint pain were observed for baricitinib versus placebo and adalimumab at week 12 ($p \leq 0.05$).

Conclusions Baricitinib provided significantly greater improvement in most PROs compared with placebo and adalimumab, including physical function MJS, pain, fatigue and quality of life. Improvement was maintained to the end of the study (week 52).

Trial registration NCT01710358.

INTRODUCTION

Rheumatoid arthritis (RA) is characterised by inflammatory activity and joint damage that often result in disability, pain, limitations in physical function and other impairments important to patients. Outcomes can be improved with effective therapy.[1–3] Decreases in physical function may be a consequence of both disease activity and irreversible, progressive joint damage.[4–6]

Patient-reported outcome (PRO) measures include health-related quality of life (HRQOL), physical function, disability, fatigue, sleep, mental health status, work productivity and work activity impairment.[7] These are standardised measures, and minimum clinically important differences (MCIDs) have been determined for many. Because these PRO measures are obtained directly from the patients, they may more accurately reflect how the patient feels and functions in relation to RA and to therapy.[8 9] The PRO measures also may facilitate doctor–patient communication and shared decision making to improve the quality of patient care.[10–12]

Baricitinib is a selective inhibitor of Janus kinase (JAK)1/JAK2 that interrupts signalling in pathways believed to be important in RA pathogenesis. The efficacy of baricitinib has been demonstrated in clinical studies in patients with RA.[13–17] In the phase 3 RA-BEAM clinical trial (NCT01710358), baricitinib 4 mg once daily (QD) was associated with clinical improvement and inhibition of progression of radiographic joint damage compared with both placebo and adalimumab in patients with RA and an inadequate response to methotrexate (MTX). Specifically, 70% of patients treated with baricitinib achieved the American College of Rheumatology 20% response rate compared with 40% of placebo-treated patients and 61% of adalimumab-treated patients. This manuscript describes the PRO data collected in the RA-BEAM clinical trial of baricitinib.[18]

METHODS

Patients and study design

RA-BEAM was a randomised, double-blind, double-dummy, placebo-controlled and active-controlled, parallel-arm, 52-week study conducted at 281 centres in 26 countries. Detailed methods of the RA-BEAM study have been published previously.[18] Briefly, patients were ≥18 years old with active RA (≥6/68 tender and ≥6/66 swollen joints; serum

high-sensitivity C reactive protein ≥6 mg/L). Patients had inadequate response to MTX. At baseline, patients were required to have either ≥3 joint erosions (based on radiographs) or >1 joint erosion with seropositivity for rheumatoid factor or anticitrullinated peptide antibodies.

Patients were randomised 3:3:2 to receive placebo, baricitinib 4 mg once daily, or biweekly subcutaneous adalimumab 40 mg, in addition to their existing background therapy (including MTX). The primary analysis time point for the study was at week 12. At week 24, patients receiving placebo switched to baricitinib. At week 16, patients whose tender and swollen joint counts improved from baseline by <20% at both weeks 14 and 16 were assigned rescue treatment (baricitinib 4 mg). After week 16, patients could be rescued at investigators' discretion based on joint counts. The study was conducted in accordance with ethical principles of the Declaration of Helsinki and Good Clinical Practice guidelines and was approved by each centre's institutional review board or ethics committee. All patients provided written informed consent.

Study endpoints and assessments

PROs were prespecified as secondary objectives of the study. Physical function was measured using the Health Assessment Questionnaire-Disability Index (HAQ-DI).[19 20] Scores range from 0 to 3, with lower scores reflecting better physical function and, thus, less disability. The HAQ-DI score changes were assessed in the context of an MCID of 0.22,[21] and the percentage of patients who reported scores that met or exceeded the population normative value (<0.5) was also assessed.[22] Disease activity and arthritis pain were measured using the Patient's Global Assessment of Disease Activity (PtGA) and the patient's assessment of pain visual analogue scales (VAS; 0–100 mm) in which higher scores indicate more disease activity or pain. Fatigue was assessed by the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) scale (range 0–52), with higher scores representing less fatigue.[23] A 3-point to 4-point change has been considered an MCID,[23–25] and in this study a value of 3.56[25] was used to assess the clinical relevance of changes in FACIT-F scores. The percentage of patients who reported scores that met or exceeded the population normative value (≥40.1) was also assessed.[23]

Some PROs were recorded using a daily electronic diary (referred to as diary PROs) from day 1 through week 12. These included duration of morning joint stiffness (MJS), and the novel MJS Severity, Worst Tiredness and Worst Joint Pain numeric rating scales (NRS). The scores for the NRS range from 0 to 10, with 10 being the worst level.

HRQOL was evaluated using the Medical Outcomes Study Short-Form-36 (SF-36; version 2, Acute),[26 27] which assesses eight domains scored from 0 to 100 that are normalised into physical component score (PCS) and mental component score (MCS). An MCID of 5 was used to assess the clinical relevance of changes in SF-36 scores.[28 29] A sensitivity analysis with an MCID of 2.5 was also evaluated. The EuroQoL 5-Dimensions (EQ-5D) Health State Profile was also used to assess HRQOL. The EQ-5D consists of two components: a descriptive system of the respondent's health and a rating of their current health state (0–100 mm VAS).[30] The UK and US scoring algorithms provide an index score using the UK or US population weighting to normalise it to that population.[31 32]

The Work Productivity and Activity Impairment Questionnaire-Rheumatoid Arthritis (WPAI-RA) was used to measure the health and symptoms of overall work productivity and impairment of regular activities, as measured during the past 7 days. Scores are calculated as impairment percentages[33] with higher percentages indicating greater impairment and less productivity.

Non-diary PROs were assessed at baseline and at weeks 1, 2, 4 and every 4 weeks thereafter to week 32 and after week 32, they were assessed at weeks 40 and 52. There were, however, some exceptions to the schedule before week 4. The FACIT-F, SF-36 and the EQ-5D were assessed at baseline and week 4, and the WPAI-RA was assessed at baseline, week 2 and week 4. These measures followed the same schedules as the other PROs after Week 4.

Statistical analyses

Randomised patients treated with ≥1 dose of placebo, baricitinib or adalimumab were included in the efficacy analyses on the basis of a modified intention-to-treat principle (analysis set).

Least squares mean (LSM) change from baseline for treatment comparisons of continuous efficacy variables were obtained using analysis of covariance with treatment, geographical region, baseline joint erosion status and baseline value in the model. For diary PRO data, analyses were based on the average of scores collected in the 7 days prior to the study visit, without baseline adjustment, until the week 12 visit date; daily scores from the day of randomisation (day 1) up to day 28 were also assessed. For the daily score analysis, mixed models for repeated measures were applied, with duration of MJS analysed by non-parametric methods.

As per the predefined analysis plan, patients who were rescued or discontinued from study or study treatment were thereafter defined as non-responders (non-responder imputation) for analysis of all categorical efficacy measures. For continuous efficacy measures, modified last observation carried forward was used, where the last observation before rescue or discontinuation was used for all subsequent time points. The WPAI-RA measures were censored after rescue or discontinuation without imputation applied.

All analyses are based on a significance level of 0.05 (two sided). p Values were not adjusted for multiple comparisons.

RESULTS

Patient disposition and baseline characteristics

A total of 1307 patients were randomised (488 placebo, 487 baricitinib 4 mg and 330 adalimumab) and 1305 patients received treatment. Patient disposition has been previously reported.[18] Baseline patient characteristics and disease activity were similar among groups (online supplementary file 1). [18] Most patients (>99%) were receiving background MTX. The majority of patients had received ≥2 prior conventional synthetic disease modifying antirheumatic drugs (csDMARDs). Baseline PROs indicated a significant disease burden, consistent with the baseline clinical disease activity (online supplementary file 1).

Patient-reported outcomes

HAQ-DI, PtGA and pain

As reported in Taylor, *et al*,[18] for HAQ-DI, PtGA, and the patient's assessment of pain, statistically significant improvements in the baricitinib group versus placebo were evident as early as week 1, the first assessment. Significant improvements in physical function and reductions in PtGA and pain were maintained at week 12 and through week 52, the end of the study (table 1). When compared with adalimumab, statistically significant improvements in HAQ-DI were seen as early as week 4 and

Table 1 Least squares mean change from baseline at 12 and 52 weeks for PRO

PRO measures (95% CI)	Week 12			Week 52	
	Placebo (n=488)	Baricitinib (n=487)	Adalimumab (n=330)	Baricitinib (n=487)	Adalimumab (n=330)
Physical function (HAQ-DI)	−0.34 (−0.39, −0.29)	−0.66*** †† (−0.71, −0.61)	−0.56*** (−0.62, −0.50)	−0.77†† (−0.83, −0.71)	−0.66 (−0.73, −0.59)
Patient's Global Assessment of Disease Activity (PtGA)	−16.7 (−18.9, −14.6)	−31.2*** †† (−33.3, −29.1)	−26.6*** (−29.1, −24.1)	−36.3††† (−38.7, −33.9)	−30.3 (−33.1, −27.5)
Patient's Assessment of Pain	−17.1 (−19.4, −14.9)	−31.5*** †† (−33.7, −29.3)	−26.4*** (−29.0, −23.7)	−36.1††† (−38.6, −33.7)	−30.3 (−33.1, −27.5)
EuroQol-5-Dimensions (EQ-5D)					
Health State Index Score, UK algorithm	0.102 (0.084, 0.119)	0.184*** (0.167, 0.202)	0.167*** (0.146, 0.188)	0.217† (0.197, 0.238)	0.182 (0.158, 0.206)
Health State Index Score, US algorithm	0.071 (0.058, 0.083)	0.130*** (0.118, 0.142)	0.117*** (0.102, 0.131)	0.154† (0.139, 0.169)	0.129 (0.112, 0.146)
VAS	7.7 (5.6, 9.8)	14.8*** †† (12.8, 16.9)	10.1 (7.7, 12.6)	19.1††† (16.6, 21.5)	11.6 (8.8, 14.5)

*p≤0.05, **p≤0.01,***p≤0.001 versus placebo.
†p≤0.05, ††p≤0.01, †††p≤0.001 versus adalimumab.
HAQ-DI, Health Assessment Questionnaire-Disability Index; PRO, patient-reported outcomes; VAS, visual analogue scale.

at week 2 for PtGA and pain, respectively; these improvements were maintained at week 12 and through week 52.

The percentages of patients who reported improvements that met or exceeded the HAQ-DI MCID of ≥0.22 at week 12 for placebo, baricitinib and adalimumab, respectively, were 58%, 75%, and 71% (p≤0.001 for baricitinib vs placebo and adalimumab vs placebo; p=0.302 for baricitinib vs adalimumab) and were 68% and 58% at week 52 for baricitinib vs adalimumab (p≤0.01). The percentage of patients who reported scores that met or exceeded the population normative value of <0.5 at week 12 or those who met or exceeded at week 52 ranged from 24% to 32% for baricitinib and adalimumab (online supplementary file 2). The percentages for baricitinib and adalimumab were statistically different (p<0.05) than placebo at week 12.

MJS duration, MJS severity, worst tiredness and worst joint pain

Baricitinib treatment resulted in significant improvement versus placebo and adalimumab for all four measures at the primary time point of the study, week 12 (table 2). Improvements versus placebo were significant from week 1 for severity of MJS, Worst Tiredness and Worst Joint Pain and from week 2 for the duration of MJS. Improvements of baricitinib versus adalimumab were observed from as early as week 2 for Worst Joint Pain, week 4 for severity of MJS and week 8 for Worst Tiredness; for duration of MJS, improvements were statistically different at weeks 1 and 12, as reported by Taylor *et al*.[18]

Consistent with the weekly averaged data, the daily diary scores showed significant improvement in patients receiving baricitinib compared with both placebo and adalimumab. Improvements relative to placebo were observed as early as day three for the severity of MJS, Worst Tiredness, and Worst Joint Pain and by day five for the duration of MJS (figure 1). Improvements relative to adalimumab were observed as early as day 19 for the severity of MJS, day 21 for Worst Tiredness and day 17 for Worst Joint Pain. No significant differences were observed in the first 28 days between baricitinib and adalimumab for the duration of MJS.

FACIT-F

Treatment with baricitinib or adalimumab was associated with significant improvements in FACIT-F at the first assessment of the measure at week 4 (p≤0.001 for baricitinib

Table 2 Day 1, week 1 and week 12 data from patient daily diaries

PRO measures	Day 1 median (IQR)			Week 1 median (IQR)			Week 12 median (IQR)		
	Placebo (n=488)	Baricitinib (n=487)	Adalimumab (n=330)	Placebo (n=488)	Baricitinib (n=487)	Adalimumab (n=330)	Placebo (n=488)	Baricitinib (n=487)	Adalimumab (n=330)
Duration of morning joint stiffness, minutes, median (IQR)	60.0 (30.0, 180.0)	60.0 (30.0, 180.0)	60.0 (20.0, 180.0)	87.5 (32.5, 180.0)	75.0† (27.5, 154.3)	60.0*** (20.0, 150.0)	60.0 (17.1, 154.3)	27.1*** † (4.3, 90.0)	36.6*** (9.2, 120.0)
	Day 1 (mean (SD))			Week 1 LSM (95% CI)			Week 12 LSM (95% CI)		
Severity of morning joint stiffness	5.5 (2.2)	5.4 (2.2)	5.3 (2.3)	5.3 (5.1 to 5.4)	4.8*** (4.6 to 4.9)	4.7*** (4.5 to 4.8)	4.1 (3.9 to 4.3)	3.0*** †† (2.8 to 3.2)	3.5*** (3.2 to 3.7)
Worst Tiredness	5.6 (2.2)	5.6 (2.2)	5.5 (2.2)	5.3 (5.1 to 5.4)	4.9*** (4.7 to 5.0)	4.8*** (4.7 to 5.0)	4.3 (4.1 to 4.5)	3.6 *** † (3.4 to 3.8)	3.9** (3.6 to 4.1)
Worst Joint Pain	5.9 (2.1)	5.9 (2.1)	5.7 (2.2)	5.6 (5.5 to 5.8)	5.0*** (4.9 to 5.2)	5.1*** (4.9 to 5.2)	4.6 (4.4 to 4.8)	3.4*** ††† (3.2 to 3.6)	4.0*** (3.8 to 4.2)

*p≤0.05; **p≤0.01; ***p≤0.001 versus placebo.
†p≤0.05, ††p≤0.01; †††p≤0.001 versus adalimumab.
LSM, least squares mean; PRO, patient-reported outcome.

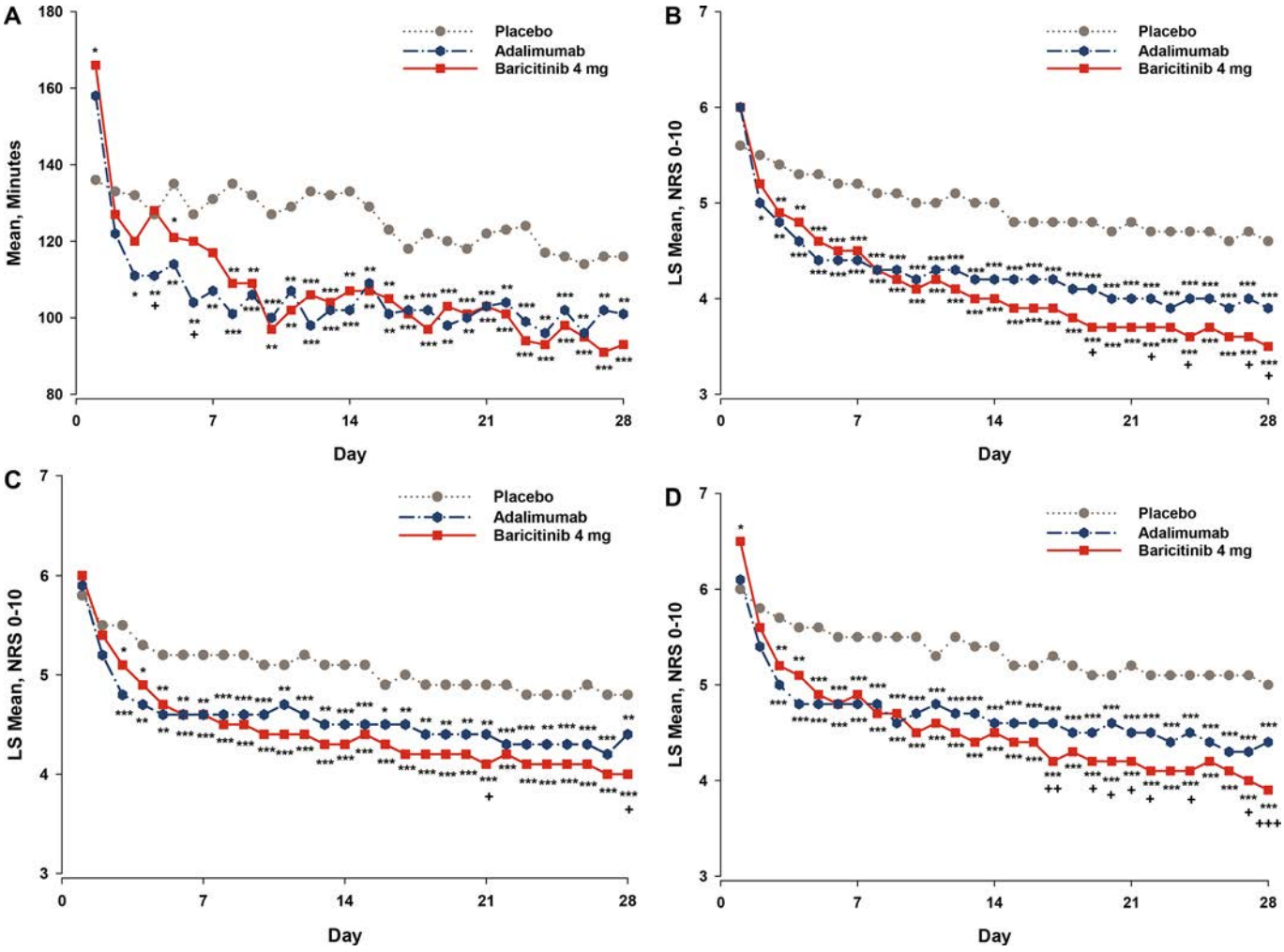


Figure 1 Change from baseline over time for the patient-reported ouctomes collected by the daily patient electronic diaries (data and either SD or 95% CIs are presented in online supplementary file 5). (A) Duration of Morning Joint Stiffness: data are mean duration of morning joint stiffness in minutes, based on daily diary entries. Daily question: 'Please indicate how long your morning joint stiffness lasted today'. Indications of statistical significance based on analysis of median difference. (B) Severity of Morning Joint Stiffness: data are LS mean scores for severity of morning joint stiffness, based on daily diary entries. Higher values indicate greater severity with numeric rating scale anchors (0–10). Daily question: 'Please rate the overall level of morning joint stiffness you had from the time you woke up today'. (C) Worst Tiredness: data are LS mean scores for Worst Tiredness, based on daily diary entries. Higher values indicate greater tiredness with numeric rating scale anchors (0–10). Daily question: 'Please rate your tiredness by selecting the one number that describes your tiredness at its worst in the last 24 hours'. (D) Worst Joint Pain: data are LS mean scores for Worst Joint Pain, based on daily diary entries. Higher values indicate greater pain with numeric rating scale anchors (0–10). Daily question: 'Please rate your joint pain by selecting the one number that describes your joint pain at its worst in the last 24 hours'. LS, least squares; NRS, numeric rating scale.p Value versus placebo: *p≤0.05; **p≤0.01; ***p≤0.001.p Value versus adalimumab: + p≤0.05; ++ p≤0.01; +++ p≤0.001.

vs placebo; p≤0.01 for adalimumab vs placebo; figure 2). The improvements in the FACIT-F score were sustained to week 24 for both baricitinib and adalimumab versus placebo (p≤0.001) and were significant at weeks 20, 28 and 52 for baricitinib versus adalimumab (p≤0.05; figure 2).

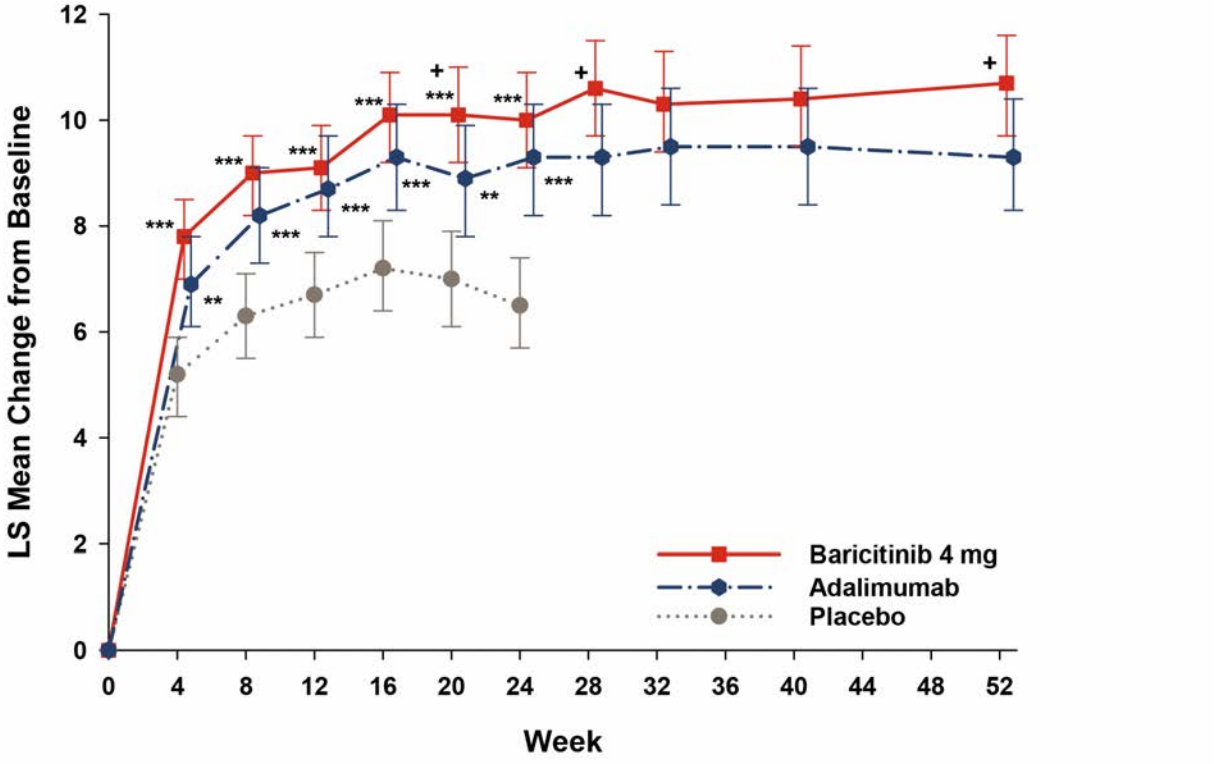
For the FACIT-F, the percentage of patients who reported improvements that met or exceeded the MCID (≥3.56) was 59%, 66% and 68% for placebo, baricitinib and adalimumab, respectively (p≤0.05 for baricitinib vs placebo; p≤0.01 for adalimumab vs placebo), at week 12 and were 60% and 54% at week 52 for baricitinib and adalimumab, respectively (p=0.084; figure 2). The percentage of patients who reported scores that met or exceeded the population normative value of ≥40.1 ranged from 41% to 46% for baricitinib and adalimumab at weeks 12 and 52 (online supplementary file 2).

Health-related quality of life

SF-36

Patients treated with baricitinib or adalimumab showed statistically significantly improved differences compared with placebo in most of the eight SF-36 domains at week 12 except for the mental health domain (both baricitinib and adalimumab) and role-emotional domain (adalimumab), which improved but did not achieve statistical significance (table 3). Compared with adalimumab, patients treated with baricitinib showed statistically significant improvement in most of the domains at week 52, except for the mental health domain (table 3).

Compared with placebo, SF-36 PCS was statistically significantly improved for patients treated with baricitinib and adalimumab (figure 3A) from the first postbaseline assessment at week 4 and was maintained through weeks 12 and 52. At week 12, the percentage of patients who met or exceeded the MCID (≥5) for



	% of patients who met or exceeded the MCID ≥ 3.56									
Week	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**40**	**52**
Placebo (N=488)	53	56	59	56	52	43	---	---	---	---
Baricitinib 4 mg (N=487)	63**	67***	66*	69***	68***	65***	64+	62	62+	60
Adalimumab (N=330)	59	66**	68**	65*	62**	59***	57	57	55	54

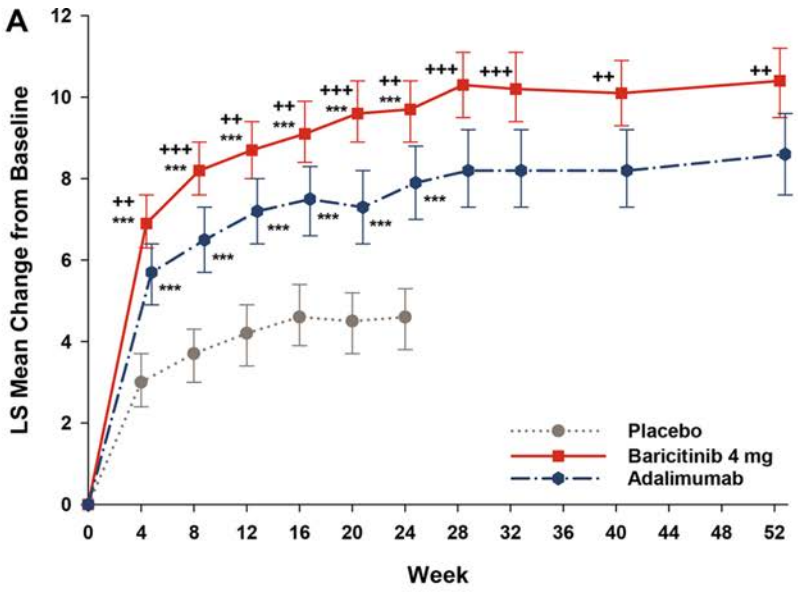
Figure 2 Change from baseline over time for the FACIT-F. Higher scores indicate less fatigue. Range=0–52. FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; MCID, minimum clinically important differences.p Value versus placebo: *p≤0.05; **p≤0.01; ***p≤0.001.p Value versus adalimumab: +p≤0.05; ++p≤0.01; +++p≤0.001.

placebo, baricitinib and adalimumab, respectively, was 40%, 65% and 56% at week 12 (for both groups vs placebo, p≤0.001 and baricitinib vs adalimumab, p≤0.05) and was 60% and 52% at week 52 for baricitinib versus adalimumab (p≤0.05). For the SF-36 MCS measure, numeric, but not statistically significant differences in the change from baseline were found for both baricitinib and adalimumab versus placebo at all time points, except for baricitinib versus placebo at week 24 (p≤0.01). The

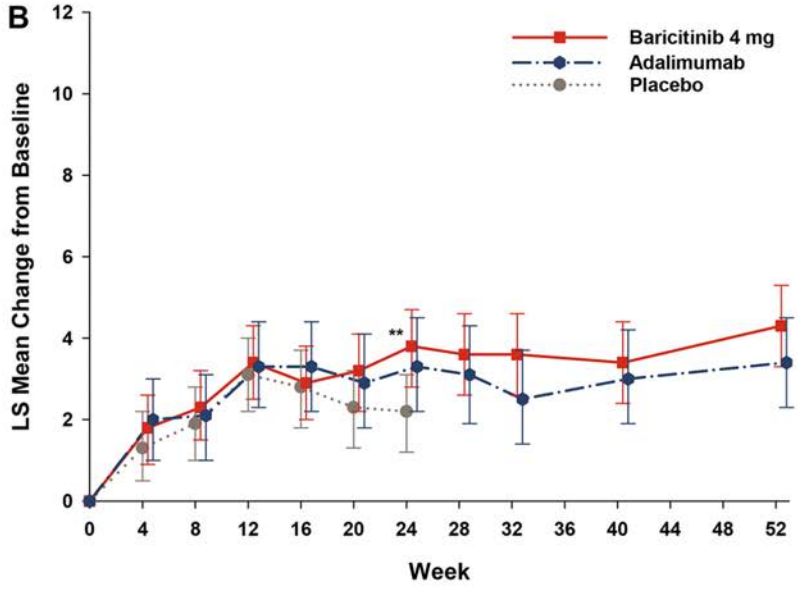
Table 3 Baseline values and least squares mean changes from baseline at weeks 12 and 52 for SF-36 domain scores.

		Baseline mean (SD)			Week 12 LSM		Week 52 LSM	
SF-36 domain scores	**Placebo (n=488)**	**Baricitinib (n=487)**	**Adalimumab (n=330)**	**Placebo (n=488)**	**Baricitinib (n=487)**	**Adalimumab (n=330)**	**Baricitinib (n=487)**	**Adalimumab (n=330)**
Physical functioning	32.4 (10.4)	32.3 (10.2)	31.6 (10.7)	4.3	8.0***†	6.8***	9.9†	8.4
Role physical	36.3 (10.3)	35.5 (10.3)	34.5 (10.5)	4.4	7.8***	6.7***	9.4††	7.5
Bodily pain	34.9 (7.7)	34.6 (7.5)	34.5 (8.5)	4.6	9.1***†	7.6***	11.2†	9.7
General health	36.6 (8.6)	37.3 (8.1)	36.3 (8.7)	3.1	5.4***	4.5†††	6.1†	4.8
Vitality	43.9 (10.1)	43.8 (9.5)	43.2 (10.5)	3.9	6.4***	5.7†††	7.9†	6.6
Social functioning	41.3 (11.3)	40.9 (11.6)	40.0 (12.2)	3.0	5.6***	4.4*	6.6††	4.6
Role emotional	41.4 (12.5)	41.4 (12.5)	40.3 (12.9)	3.7	5.1*	4.8	6.6†	5.3
Mental health	42.9 (11.3)	43.3 (11.1)	42.5 (11.5)	3.7	4.0	3.9	5.1	4.4

*p≤0.05, **p≤0.01, ***p≤0.001 versus placebo.
†p≤0.05,††p≤0.01,†††p≤0.001 versus adalimumab.
LSM, least squares mean; SF-36, Short-Form-36.



	% of patients who met or exceeded the MCID ≥ 5									
Week	4	8	12	16	20	24	28	32	40	52
Placebo (N=488)	34	39	40	41	40	35	---	---	---	---
Baricitinib 4 mg (N=487)	56***	64***++	65***+	67***++	68***++	65***+	66+++	61+	59	60+
Adalimumab (N=330)	51***	54***	56***	56***	55***	57***	54	53	53	52



	% of patients who met or exceeded the MCID ≥ 5									
Week	4	8	12	16	20	24	28	32	40	52
Placebo (N=488)	31	32	33	32	27	28	---	---	---	---
Baricitinib 4 mg (N=487)	33	33	37	36	36**	38**	34	36	32	34
Adalimumab (N=330)	32	34	37	36	34*	34	33	32	32	29

Figure 3 Change from baseline for the physical and mental component score for the SF-36. (A) Physical component score: data in table are % patients who met or exceeded the minimum clinically important difference in SF-36 PCS (≥5 points). Higher scores indicate improvement. (B) Mental component score: data in table are % patients who met or exceeded the minimum clinically important difference in SF-36 MCS (≥5 points). Higher scores indicate improvement. p Value versus placebo: *p≤0.05; **p≤0.01; ***p≤0.001.p Value versus adalimumab: +p≤0.05; ++p≤0.01; +++p≤0.001.

proportion of patients who reported improvements that met or exceeded the MCID for the MCS was not statistically significantly different from placebo for either group at any time point, except at weeks 20 (both baricitinib and adalimumab differed from placebo, p≤0.05) and at week 24 (p≤0.01 for baricitinib vs placebo; figure 3B). Results were similar for the MCID value of 2.5 (online supplementary file 3).

EQ-5D

A statistically significant improvement in the EQ-5D index scores (both US and UK) were observed at the first postbaseline assessment, week 4 (data not shown), for both baricitinib and adalimumab versus placebo and was maintained to week 12 (table 1). By week 52, statistically significant improvements in EQ-5D index scores were observed for baricitinib versus adalimumab (table 1). For the EQ-5D VAS at 4 weeks, baricitinib-treated and adalimumab-treated patients showed statistically significant improvement compared with placebo-treated patients (data not shown). By week 12, however, statistically significant improvement in EQ-5D VAS was observed for only baricitinib-treated patients (p≤0.001 vs placebo; p≤0.01 versus adalimumab); this was maintained through week 52 for baricitinib versus adalimumab (p≤0.001; table 1).

Work Productivity and Activity Impairment

At baseline, 41%–43% of the patients were employed. Patients treated with baricitinib reported statistically significantly improved daily activity compared with placebo and adalimumab at week 12 (p≤0.001 for both groups vs placebo; p≤0.01 for baricitinib vs adalimumab); improvements compared with adalimumab, however, were not statistically significant at week 52 (online supplementary file 4). Among those patients who were employed at baseline and those who maintained employment at week 12, statistically significant improvements in absenteeism (p≤0.05), presenteeism (p≤0.001) and work productivity loss (p≤0.001) were seen with baricitinib compared with placebo; improvements compared with adalimumab, however, were not statistically significant at week 52. Only work productivity loss was statistically significantly improved with baricitinib versus adalimumab at week 12 (p≤0.05; online supplementary file 3).

DISCUSSION

The RA-BEAM study evaluated baricitinib 4 mg once daily in patients with an inadequate response to MTX who were naive to biological DMARDs using placebo and adalimumab 40 mg biweekly as comparators.[18] Patients continued to take stable background csDMARDs (including MTX) during the study. This paper evaluates whether the clinical efficacy data for baricitinib were complemented by corresponding changes in PROs.

Baseline PROs describe substantial duration (≥60 min) and severity of MJS, severe impairment of physical function and high levels of pain and fatigue (including tiredness) among patients enrolled in the study. Baricitinib treatment produced significantly greater improvements compared with placebo and adalimumab in most of the prespecified PROs, including physical function, pain, fatigue, duration and severity of MJS and HRQOL at week 12. Furthermore, baricitinib produced rapid improvements in the diary PROs compared with placebo and adalimumab, with significant differences vs placebo appearing within days of initiating treatment. Improvements were maintained to week 52 compared with adalimumab in physical function, pain, fatigue and HRQOL (eg, SF-36 PCS and EQ-5D).

In this analysis and in the results presented by Taylor *et al*,[18] treatment with baricitinib resulted in a rapid improvement in PROs; patients showed statistically significant improvements as early as week 1 in HAQ-DI, PtGA and the patient's assessment of pain, and these results were maintained until the end of the trial at week 52. Compared with placebo, a significantly greater proportion of patients treated with baricitinib or adalimumab reported improvements that met or exceeded the MCID and the population normative values for HAQ-DI and FACIT-F at week 12.

Similar results were seen for duration and severity of MJS, Worst Tiredness and Worst Joint Pain as assessed using the patient daily diaries and improvement continued to week 12. The rapid onset of action, with improvement in relevant signs and symptoms (such as pain and tiredness) as early as 3 days after the start of treatment, is a useful complement to the efficacy observed at the later time points.

Consistent with these results, patients treated with baricitinib reported improvements in HRQOL, as measured by the EQ-5D and SF-36 PCS compared with placebo and adalimumab. For the SF-36, improvements across most of the SF-36 domains were observed for baricitinib and adalimumab compared with placebo at week 12 and for baricitinib compared with adalimumab at week 52. Furthermore, when a five-point change was used for the MCID, it was found that 65% of the baricitinib-treated patients met or exceeded the MCID for the SF-36 PCS (figure 3A). In contrast with the PCS, no statistically significant differences were observed between baricitinib-treated and adalimumab-treated patients compared with placebo-treated patients with the SF-36 MCS. Across treatment groups at baseline, the SF-36 MCS values ranged from 46 to 47, which are close to the population normative data of 50.[28] This suggests only modest impairment for the MCS at baseline such that a marked improvement with therapy would not be expected. This SF-36 MCS result aligns with previous results from other trials.[34–36]

Compared with placebo, patients treated with baricitinib showed statistically significant improvement across all scores of the work productivity assessment at week 12. When compared with adalimumab, the baricitinib-treated patients showed statistically significant improvement in work productivity loss and impairment of regular activity at week 12; these improvements continued through week 52 but were not statistically significantly different.

The results from this analysis are similar to those observed in other phase 3, randomised clinical trials of baricitinib in different patient populations.[34–36]

The limitations of this analysis include the use of carrying forward the last observations before rescue or discontinuation. This method assumes that the PRO values do not change over time. Also, as in most double-blind comparator trials, the inclusion and exclusion criteria restrict patient participation such that these results may not be fully generalisable to the population seen in clinical practice.

This study used well-established PRO measures that can holistically evaluate the burden of RA and the treatment effects across many health domains. Some of the PRO measures are incorporated into the patient ratings in the ACR core set, while others, such as the EQ-5D and SF-36, are established HRQOL instruments that may more broadly measure the effects of RA and treatment on patients. The use of diary records allowed patients to report the impact of symptoms of importance to them as they arose, therefore permitting a more complete evaluation than by means of periodic recording of recollected symptoms. Furthermore,

these PRO measures may help facilitate discussions between patients and their healthcare providers; they may help address patient concerns such as how long it will take to feel improvement (onset of action), how long to try the new treatment before determining that it is not effective (efficacy plateau) and how long the treatment will be effective (sustainability or the risk of relapse). In addition to facilitating the physician–patient dynamic, PRO measures are being increasingly used in randomised clinical trials and allow for epidemiological assessments across different patient populations and disease states. PRO assessments such as the work productivity measure also provide an insight into the broader, societal impact of RA.

The use of a variety of PRO measures also allows for an assessment of the clinical importance of the present study's results. Similar trends were observed between comparisons of baricitinib with both placebo and adalimumab in many of the PROs. Additionally, some PRO measures were assessed with established and validated MCID values. Statistically significant differences in MCIDs between treatment groups implies clinical significance on a group level. Collectively, the results of the present study demonstrate treatment benefits for baricitinib that appear clinically relevant.

The RA-BEAM study demonstrated that patients treated with baricitinib experienced a greater improvement compared with patients treated with placebo or adalimumab in most PROs across different domains of RA, including physical function, MJS, fatigue, pain and HRQOL. These improvements tended to occur within the first weeks of treatment and were maintained throughout the 52-week trial.

Correction notice This article has been corrected since it published Online First. The 'patient disposition and baseline characteristics' paragraph has been updated.

Acknowledgements The authors would like to thank Barbara Coffey, MS, and Molly Tomlin, MS, of Eli Lilly and Company for their assistance with manuscript preparation and process support and Scott Beattie, PhD, for statistical support.

Contributors All authors have contributed to the work and approve the presented findings.

Funding This study was funded by Eli Lilly and Company and Incyte Corporation.

Competing interests ECK has received grant/research support, consulting support or speaker bureau fees from Abbott, Amgen, AstraZeneca, Biotest, Bristol Myers Squibb, F. Hoffmann-La Roche Inc., Janssen, Eli Lilly and Company, Genentech, Janssen, Merck, Novartis, Pfizer, Sanofi-Aventis and UCB. PCT has received grant/research support or consulting support from AbbVie, Bristol Myers Squibb, Celgene, Eli Lilly and Company, Galapagos, GlaxoSmithKline, Merck, Pfizer, Takeda and UCB. YT has received grant/research support or speaker bureau fees from AbbVie, Astellas, Asahi Kasei, Bristol Myers Squibb, Chugai, Daiichi-Sankyo, Eisai, Eli Lilly and Company, GlaxoSmithKline, Janssen, Kyowa-Kirin, Mitsubishi-Tanabe, MSD, Pfizer, Santen, Takeda, Taisho-Toyama, Teijin and UCB. CG, AMD, TR, SdB, VA and BL are employees of Eli Lilly and Company, and all except AMD are shareholders of Eli Lilly and Company. AD is a consultant for Eli Lilly and Company. JVZ is a consultant for Eli Lilly and Company. JACC is a consultant for Eli Lilly and Company. MEW has received grant/research support or consulting support from AbbVie, Amgen, Bristol-Myers Squibb, Canfite, Corrona, Crescendo Bioscience, Genzyme, Idera, Infinity, Lycera, Eli Lilly and Company, Medimmune, Merck, Merck/serona, Momenta, Novartis, Pfizer, Roche, Samsung, Sanofi and UCB.

Ethics approval The study was approved by each center's institutional review board or ethics committee.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Carr A, Hewlett S, Hughes R, *et al*. Rheumatology outcomes: the patient's perspective. *J Rheumatol* 2003;30:880–3.
2 Strand V, Cohen S, Crawford B, *et al*. Patient-reported outcomes better discriminate active treatment from placebo in randomized controlled trials in rheumatoid arthritis. *Rheumatology* 2004;43:640–7.
3 Klarenbeek NB, Güler-Yüksel M, van der Kooij SM, *et al*. The impact of four dynamic, goal-steered treatment strategies on the 5-year outcomes of rheumatoid arthritis patients in the BeSt study. *Ann Rheum Dis* 2011;70:1039–46.
4 Drossaers-Bakker KW, de Buck M, van Zeben D, *et al*. Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. *Arthritis Rheum* 1999;42:1854–60.
5 Aletaha D, Smolen J, Ward MM. Measuring function in rheumatoid arthritis: identifying reversible and irreversible components. *Arthritis Rheum* 2006;54:2784–92.
6 Smolen JS, Aletaha D, Grisar JC, *et al*. Estimation of a numerical value for joint damage-related physical disability in rheumatoid arthritis clinical trials. *Ann Rheum Dis* 2010;69:1058–64.
7 Kekow J, Moots RJ, Emery P, *et al*. Patient-reported outcomes improve with etanercept plus methotrexate in active early rheumatoid arthritis and the improvement is strongly associated with remission: the COMET trial. *Ann Rheum Dis* 2010;69:222–5.
8 Gossec L, Dougados M, Dixon W. Patient-reported outcomes as end points in clinical trials in rheumatoid arthritis. *RMD Open* 2015;1:e000019.
9 Guidance for industry: patient-reported outcome measures: use in Medical Product Development to support labeling claims. Secondary Guidance for Industry: Patient-Reported Outcome Measures Use in Medical Product Development to Support Labeling Claims. 2009 http://www.fda.gov/downloads/Drugs/./Guidances/UCM193282.pdf (accessed 17 Apr 2017).
10 Fries JF, Spitz P, Kraines RG, *et al*. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
11 Sokka T. Morning stiffness and other patient-reported outcomes of rheumatoid arthritis in clinical practice. *Scand J Rheumatol Suppl* 2011;125:23–7.
12 Smolen JS, Landewe R, Bijlsma J, *et al*. EULAR recommendations for themanagement of rheumatoid arthritis with synthetic and biologicaldisease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017:1–18.
13 Keystone EC, Taylor PC, Drescher E, *et al*. Safety and efficacy of baricitinib at 24 weeks in patients with rheumatoid arthritis who have had an inadequate response to methotrexate. *Ann Rheum Dis* 2015;74:333–40.
14 Tanaka Y, Emoto K, Cai Z, *et al*. Efficacy and safety of Baricitinib in Japanese patients with active rheumatoid Arthritis Receiving Background Methotrexate therapy: a 12-week, Double-blind, randomized Placebo-controlled study. *J Rheumatol* 2016;43:504–11.
15 Fleischmann R, Schiff M, van der Heijde D, *et al*. Baricitinib, Methotrexate, or combination in patients with Rheumatoid Arthritis and no or Limited Prior Disease-Modifying Antirheumatic Drug Treatment. *Arthritis Rheumatol* 2017;69:506–17.
16 Dougados M, van der Heijde D, Chen YC, *et al*. Baricitinib in patients with inadequate response or intolerance to conventional synthetic DMARDs: results from the RA-BUILD study. *Ann Rheum Dis* 2017;76:88–95.
17 Genovese MC, Kremer J, Zamani O, *et al*. Baricitinib in patients with refractory rheumatoid Arthritis. *N Engl J Med* 2016;374:1243–52.
18 Taylor PC, Keystone EC, van der Heijde D, *et al*. Baricitinib versus Placebo or Adalimumab in Rheumatoid Arthritis. *N Engl J Med* 2017;376:652–62.
19 Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23(Suppl 39):S14–18.
20 Ramey D, Fries J, Singh G. The Health Assessment Questionnaire 1995: status and review. Spiker B, ed. *Quality of life and pharmacoeconomics in clinical trials*. 2nd ed. Philadelphia: Lippincott-Raven, 1996:227–37.
21 Wells GA, Tugwell P, Kraag GR, *et al*. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. *J Rheumatol* 1993;20:557–60.
22 Navarro-Compán V, Smolen JS, Huizinga TW, *et al*. Quality indicators in rheumatoid arthritis: results from the METEOR database. *Rheumatology* 2015;54:1630–9.
23 Cella D, Yount S, Sorensen M, *et al*. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. *J Rheumatol* 2005;32:811–9.
24 Strand V, Burmester GR, Zerbini CA, *et al*. Tofacitinib with methotrexate in third-line treatment of patients with active rheumatoid arthritis: patient-reported outcomes from a phase III trial. *Arthritis Care Res* 2015;67:475–83.
25 Keystone E, Burmester GR, Furie R, *et al*. Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. *Arthritis Rheum* 2008;59:785–93.
26 Brazier JE, Harper R, Jones NM, *et al*. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160–4.
27 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
28 Kosinski M, Zhao SZ, Dedhiya S, *et al*. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum* 2000;43:1478–87.

29 Strand V, Singh JA. Newer biological agents in rheumatoid arthritis: impact on health-related quality of life and productivity. *Drugs* 2010;70:121–45.
30 EuroQol Group. EQ-5D-5L User Guide. Version 1.0. 2011. http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L.pdf (accessed 1 Jul 2012); http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-5L_UserGuide_2015.pdf (accessed 17 Apr 2017).
31 Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37:53–72.
32 Herdman M, Gudex C, Lloyd A, *et al*. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
33 Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics* 1993;4:353–65.
34 Smolen JS, Kremer JM, Gaich CL, *et al*. Patient-reported outcomes from a randomised phase III study of baricitinib in patients with rheumatoid arthritis and an inadequate response to biological agents (RA-BEACON). *Ann Rheum Dis* 2017;76:694–700.
35 Schiff M, Takeuchi T, Gaich C, *et al*. THU0623 Patient-Reported Outcomes from A Phase 3 Study of Baricitinib in Patients with Early Rheumatoid Arthritis Who Had Received Limited or No Treatment with Disease-Modifying anti-Rheumatic Drugs: Table 1. *Ann Rheum Dis* 2016;75263:419.1–419.
36 Emery P, Blanco R, Maldonado Cocco J, *et al*. Patient-reported outcomes from a phase III study of baricitinib in patients with conventional synthetic DMARD-refractory rheumatoid arthritis. *RMD Open* 2017;3:e000410.

EXTENDED REPORT

Subgroup analyses of the effectiveness of oral glucosamine for knee and hip osteoarthritis: a systematic review and individual patient data meta-analysis from the OA trial bank

Jos Runhaar,[1] Rianne M Rozendaal,[1] Marienke van Middelkoop,[1] Hans J W Bijlsma,[2] Michael Doherty,[3] Krysia S Dziedzic,[4] L Stefan Lohmander,[5] Timothy McAlindon,[6] Weiya Zhang,[3] Sita Bierma Zeinstra[7]

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-2017-211149).

For numbered affiliations see end of article.

Correspondence to
Jos Runhaar, Department of General Practice, Erasmus University Medical Center Rotterdam, Erasmus MC, Rotterdam Zuid-Holland, 3000 CA, The Netherlands;
j.runhaar@erasmusmc.nl

Received 16 January 2017
Revised 22 June 2017
Accepted 26 June 2017
Published Online First 28 July 2017

ABSTRACT

Objective To evaluate the effectiveness of oral glucosamine in subgroups of people with hip or knee osteoarthritis (OA) based on baseline pain severity, body mass index (BMI), sex, structural abnormalities and presence of inflammation using individual patient data.

Methods After a systematic search of the literature and clinical trial registries, all randomised controlled trials (RCTs) evaluating the effect of any oral glucosamine substance in patients with clinically or radiographically defined hip or knee OA were contacted. As a minimum, pain, age, sex and BMI at baseline and pain as an outcome measure needed to be assessed.

Results Of 21 eligible studies, six (n=1663) shared their trial data with the OA Trial Bank. Five trials (all independent of industry, n=1625) compared glucosamine with placebo, representing 55% of the total number of participants in all published placebo-controlled RCTs. Glucosamine was no better than placebo for pain or function at short (3 months) and long-term (24 months) follow-up. Glucosamine was also no better than placebo among the predefined subgroups. Stratification for knee OA and type of glucosamine did not alter these results.

Conclusions Although proposed and debated for several years, open trial data are not widely made available for studies of glucosamine for OA, especially those sponsored by industry. Currently, there is no good evidence to support the use of glucosamine for hip or knee OA and an absence of evidence to support specific consideration of glucosamine for any clinically relevant OA subgroup according to baseline pain severity, BMI, sex, structural abnormalities or presence of inflammation.

INTRODUCTION

Oral glucosamine has long been recommended for the treatment of knee and hip osteoarthritis (OA). However, recent guidelines by Osteoarthritis Research Society International (OARSI)[1] and The National Institute for Health and Care Excellence (NICE)[2] highlight the lack of support for the efficacy of oral glucosamine for the management of symptoms or disease modification in OA.[3] With increasing study quality over the past decades, reported effect sizes for glucosamine have decreased.[4] Furthermore, methodological issues in trials studying the effect of glucosamine for OA symptoms, such as inadequate allocation concealment and absence of intention-to-treat analyses, has resulted in overestimation of its effectiveness.[5] A network meta-analysis from seven high-quality, large (>200 participants per trial) randomised controlled trials (RCTs) concluded that oral glucosamine was not superior to placebo in reducing OA pain or reduction in joint space narrowing.[3]

Notwithstanding the overall lack of efficacy of glucosamine, it is possible that certain subgroups of OA might respond differently (either better or worse) to any specific treatment.[6] These subgroups might be based on different pathologies underlying the clinical presentation of OA, different disease stages or on the presence of different comorbidities.[6] Accordingly, clinical guidelines increasingly call for the identification of any predictors of response to different treatment modalities.[7] Since the effectiveness of glucosamine varies among different populations,[4 5 8] it is possible that glucosamine might show higher efficacy when targeted at specific subgroups.

Recently, van Middelkoop *et al*[9] reported on the methodology and legal structure to perform individual patient data (IPD) meta-analyses to identify clinically relevant subgroups that may show differential response to different OA treatments (the OA Trial Bank). The proposed methodologically robust method tests subgroup–treatment interaction effects using IPD from multiple published trials and allows for adjustment for confounding at both study and individual patient levels.[9] Using this method, increased short-term efficacy for glucocorticoid treatment among knee OA patients with more severe pain has been demonstrated.[10]

The present study aimed to collect IPD of all RCTs performed for oral glucosamine in people with knee and hip OA to evaluate the efficacy within predefined subgroups of OA based on pain severity, body mass index (BMI), sex, structural abnormalities and presence of inflammation.

METHODS

Systematic search

To identify all available RCTs, a systematic search of the literature was performed in PubMed, the Cochrane Central Register of Controlled Trials,

To cite: Runhaar J, Rozendaal RM, van Middelkoop M, *et al*. *Ann Rheum Dis* 2017;**76**:1862–1869.

Embase, Web of Science, Cinahl and Scopus. The search strategy was based on the search protocol of the Cochrane publication on the effectiveness of glucosamine.[8] It was adjusted for the different databases and limited to publications from 1994 because of the likelihood of communicating with corresponding authors and data being available (searched up to March 2014 and available on request). Reference lists were hand searched for further identification of published work. Additional potential ongoing studies were searched for in clinical trial registries.

Two authors (JR and RR) independently selected citations based on titles and abstracts. Subsequently, full articles were obtained for those citations thought to fulfil the inclusion criteria and were independently assessed by the two review authors. A third review author was consulted if consensus was not reached (MvM). No protocol was registered for the current project, but full protocol details for the systematic review and the IPD meta-analysis were prespecified in the data delivery licence agreement that was approved by all members of the OA Trial Bank Steering Committee before the systematic search of the literature was initiated (available on request).

Inclusion/exclusion criteria

All RCTs evaluating the effect of any oral glucosamine substance in participants with knee or hip OA were included. This included studies testing the effects of glucosamine within a subgroup of participants with OA. Studies solely testing a combination of glucosamine with another substance (eg, chondroitin) were not included. There was no language restriction.

Participants

Participants were men and/or women with a diagnosis of OA of the knee or hip:

1. according to ACR classification criteria[11] or
2. on the basis of detailed clinical and/or radiographic information.

Studies including a subgroup of knee or hip OA patients were also included, because IPD were collected.

Interventions

All comparisons between different oral glucosamine doses or between different frequencies of intake were included. Cointerventions were allowed as long as they were identically applied to the glucosamine and control group.

Comparator

All comparisons between oral glucosamine and any placebo/medication/dietary supplement/other non-surgical treatment were included.

Outcomes

The minimum criterion for inclusion of RCTs was adequate reporting of pain as an outcome measure.

Baseline predictors

1. Important data

As a minimum, severity of pain, age, sex and BMI should have been assessed at baseline in order to define subgroups.

2. If available

Signs of inflammation, either by physical examination (warmth and effusion) or by additional testing (ultrasound, MRI, biopsy and serum c-reactive protein (CRP)/erythrocyte sedimentation rate (ESR)), and structural abnormalities by radiography or magnetic resonance imaging (MRI)at baseline.

Data collection, transfer and checks

All corresponding authors of eligible trials were approached and asked to share trial data (first by email, subsequently by telephone). When corresponding authors could not be reached, the other listed authors and the institutes in which the trials had been performed were contacted. All data-deliverers willing to participate (ie, the research institutes who own the data) were asked to sign the data delivery licence agreement, including items on input data, obligations, ownership of data, terms, authorship, all subgroup analyses and publications. All anonymous data were transferred to a secured database at the Erasmus University Medical Center Rotterdam. On receiving the data, a thorough check of the data took place by reproducing the main baseline characteristics and the reported changes over time for the available outcome measures. Uncertainties were resolved in collaboration with the trialists.

Risk of bias assessment

The methodological quality of all included trials in the OA Trial Bank were assessed using the 12 criteria recommended by Cochrane (see online supplementary file 1) and were evaluated independently by two researchers (JR and RR). The criteria were scored as 'yes' (low risk of bias), 'no' (high risk of bias) or 'unclear'. Any disagreement between the review authors was resolved by discussion, including input from a third review author (MvM). A study with a low risk of bias was defined as fulfilling six or more of the criteria items. In case the number of shared studies would allow proper interpretation (≥10 studies), funnel plots were considered for evaluation of publication bias.

Data analyses

First, heterogeneity of the eligible studies was determined for the primary outcomes, using a two-stage meta-analysis approach in Review Manager V.5.3. In case of high heterogeneity (I^2 index >50), sensitivity analyses without data from trials causing the heterogeneity were planned. Second, a descriptive comparison between studies was performed. We assumed missing data to be missing at random. Therefore, missing data for covariates and outcome measures were imputed, using multiple imputation methods, within each original study. Outcomes measured on different scales were standardised in order to pool the data. Predefined subgroup factors were dichotomised, based on consensus of the OA Trial Bank Steering Committee. For this, descriptive statistics of the subgroup variables for each of the five trials were shared with the Steering Committee, together with proposed cut-off values, based on literature, data separation in the available trials and previous IPD meta-analysis by the OA Trial Bank.[10]

The primary outcome measures were pain severity in the short-term (3–6 months) and at long-term (≥1 year) follow-up. Secondary outcomes were physical function and all forms of structural changes at these time points.

A one-stage multilevel regression analysis was performed to estimate the magnitude of the effect (estimated pooled mean differences) of glucosamine over the control intervention over all included studies and in the different subgroups with the individuals nested within each study. A single covariate was added to the regression models to indicate the study (fixed factor), in order to adjust for possible residual confounding by study differences. To assess possible subgroup effects, a random-effects linear regression model was used to determine interaction effects. This model included the dependent variable (primary or secondary outcome measure), the independent variable

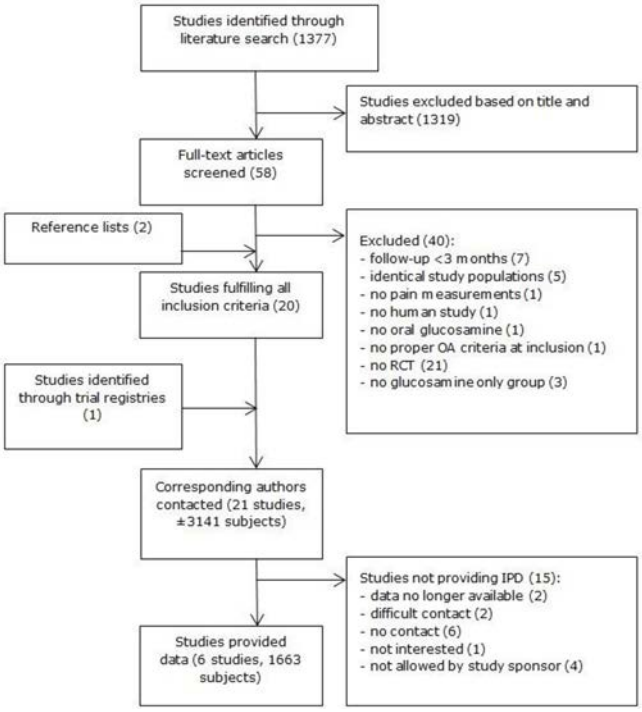


Figure 1 Study flow chart. IPD, individual patient data; OA, osteoarthritis; RCT, randomised controlled trials.

(treatment group), the effect modifier (subgroup indicator) and an interaction term (independent variable × effect modifier). All analyses were adjusted for age sex, BMI, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain at baseline and were performed with and without stratification for type of glucosamine and with and without stratification for the affected joint. Comparisons and subgroup analysis for which only one RCT was available were not taken into account, since main effects were already studied in the original publication, and individual trials usually were not powered for subgroup analysis. A p value <0.05 was regarded as statistically significant in all analyses, using IBM SPSS software V.22.

RESULTS

The literature search resulted in 1377 abstracts. After screening, 58 publications were evaluated in full-text and 18 fulfilled all inclusion criteria,[12–29] with two additional trials identified from the references of the included trials[30 31] (figure 1). Searching the clinical trial registries resulted in one additional potentially eligible trial (NCT01074476). All 21 corresponding authors of these trials were contacted for participation. After multiple efforts to contact all data owners of the eligible trials, authors/institutes of six studies agreed to participate and delivered trial data to the OA Trial Bank.[14–16 24 28 29] Corresponding authors of two trials indicated that trial data were no longer available.[13 23] Two corresponding authors did reply positively to the initial request for data sharing, but a signed licence agreement was never received.[12 20] One corresponding author was not interested in participation.[17] No contact was established with any of the authors nor the research institutes of five studies,[18 21 26 30 31] and the one study identified in the clinical trial registry. Four data owners indicated that they were not permitted to share their data by the study sponsor.[19 22 25 27] See table 1 for full details of all eligible studies.

Five out of the six studies willing to participate involved knee OA participants,[14–16 24 29] while only one involved hip OA participants.[28] Follow-up duration in the six trials ranged from 3 to 24 months. Three studies evaluated glucosamine sulfate (GS)[15 16 28] and two glucosamine hydrochloride (GH).[14 29] The publication of the remaining study stated that the first 163 subjects received GS but that the subsequent subjects received GH.[24] However, after extensive communication with the trial owner, the order of glucosamine type was deemed to be a typographical error, since the supplier of the glucosamine for the latter part of the participants (Rottapharm) is renowned for its GS. Data on participants within this trial were allocated to the stratified analysis for glucosamine type based on this new insight of the glucosamine type provided. With the exception of the trial by Coulson *et al* that used green-lipped mussel extract as comparison,[15] all studies compared their glucosamine substrate against placebo. The trial by Coulson *et al* was therefore not included in the subgroup analysis (mean change in WOMAC pain −1.6 (−3.7 to 0.6) on a 0–20 scale in favour of glucosamine (p=0.157)).[15] The trial by Sawitzke *et al*[29] presented long-term follow-up from the Clegg *et al*,[14] but since both publications report on different outcome measures of interest (clinical data and radiography vs clinical data only) and risk of bias could be assessed for both publications separately, both were indicated as separate trials. No important issues were identified when checking shared trial data, but for the trial by McAlindon[24] for which data of the first 199 (out of 205 in the original publication) could be retrieved by the trial owners. No relevant differences in baseline characteristics for the subjects with shared data and the published data were observed. Percentages of missing data for the main baseline characteristics and all outcome measures for each of the five individual trials are presented in online supplementary table 2 . All listed variables were used in the multiple imputation by the SPSS software package, creating 20 imputed data sets for each trial.

The five trials included in the analysis included a total number of 1625 participants (64% women), 815 randomised to glucosamine and 810 to placebo. This reflected 55% of the participants randomised in the 17 published RCTs on glucosamine versus placebo. Pain was measured in all five studies using the ordinal WOMAC questionnaire.[32] Scores were rescaled to a 0–100 scale and defined at short-term (closest to a minimal of 3 months follow-up) for the trials by McAlindon *et al*.,[24] Clegg *et al*,[14] and Rozendaal *et al*[28] and long-term (2 years follow-up) for Fransen *et al*.,[16] Sawitzke *et al*,[29] and Rozendaal *et al*.[28] Physical function was also measured in all five studies using the WOMAC questionnaire and was rescaled and defined in an identical matter. Figure 2 presents the overall mean differences of these five trials for the primary outcome at short term and long term, based on the imputed data sets.

The following subgroups were defined: WOMAC pain <70 versus ≥70, BMI <27 kg/m^2 versus ≥27 kg/m^2, Kellgren and Lawrence grade[33] (KL-grade) 0–2 versus KL3-4 and presence versus absence of inflammation. Presence of inflammation was defined as either presence of swelling/effusion on clinical examination[14 29] or an elevated erythrocyte sedimentation rate (ESR),[28] defined as ESR ≥20 mm/h for men aged ≥50 years, ESR ≥15 mm/h for men aged <50 years, ESR ≥30 mm/h for women aged ≥50 years and ESR ≥20 mm/h for women aged <50 years. Inflammation data were only available when combining data from one knee OA[14 29] and one hip OA trial.[28] Therefore, no additional stratification was possible. Baseline Kellgren and Lawrence grades were only available in one knee OA trial with short-term outcomes,[14] two knee OA trials with

Table 1 Characteristics of all eligible and contacted studies (stratified for authors' reply on data sharing request)

	Origin	Participants	N in control group	N in glucosamine group	Interventions	Follow-up	Funding source	Reply to data sharing request
Clegg *et al*[14]	USA	Knee OA	313	317	GH versus CS versus GH+CS versus placebo versus celecoxib	6 months	Funding agency	Data delivered to OA Trial Bank
Coulson *et al*[15]	Australia	Knee OA	21	17	GS versus green-lipped mussel extract	3 months	Commercial party	Data delivered to OA Trial Bank
Fransen *et al*[16]	Australia	Knee OA	151	152	GS versus GS+CS versus CS versus placebo	24 months	Governmental institution and by some supplementary funding from a commercial party	Data delivered to OA Trial Bank
McAlindon *et al*[24]	USA	Knee OA	104	101	GH versus placebo***	3 months	Funding agency	Data delivered to OA Trial Bank
Rozendaal *et al*[28]	The Netherlands	Hip OA	111	111	GS versus placebo	24 months	Governmental institution	Data delivered to OA Trial Bank
Sawitzke *et al*[29]	USA	Knee OA	131	134	GH versus CS versus GH+CS versus placebo verus celecoxib	24 months	Governmental institution	Data delivered to OA Trial Bank
Cibere *et al*[13]	Canada	Knee OA	66	71	GS versus placebo	6 months	Funding agency	Data no longer available
Martí-Bonmatí *et al*[23]	Spain	Knee OA	4	7	GS versus acetaminophen	6 months	Commercial party	Data no longer available
Chopra *et al*[12]	India	Knee OA	35	35	Five herbal groups versus GS versus placebo	4 months	Governmental institution	Positive to first request, but no data delivery
Hughes and Carr[20]	UK	Knee OA	40	40	GS versus placebo	6 months	Unknown	Positive to first request, but no data delivery
Frestedt *et al*[17]	USA	Knee OA	16	19	GS versus placebo versus Aquamin versus Aquamin+GS	3 months	Commercial party	Not interested in participation
Giordano *et al*[31]	Italy	Knee OA	30	30	GS versus placebo	3 months	Unknown	No contact with authors/institutions
Hatano *et al*[18]	Japan	Knee OA	31	36	Soymilk with versus without *N*-acetyl glucosamine	3 months	Unknown	No contact with authors/institutions
Kawasaki *et al*[21]	Japan	Knee OA	42	49	Home exercise versus home exercise+GH versus home exercise+risedronate	18 months	Unknown	No contact with authors/institutions
NCT01074476*	Canada	Knee OA	10	10	GS versus placebo	3 months	Governmental institution	No contact with authors/institutions
Petersen *et al*[26]	Denmark	Knee OA	12	12	GS versus placebo versus ibuprofen	3 months	Governmental institution, and funding agency	No contact with authors/institutions
Usha and Naidu[30]	India	knee OA	28	30	G vs. MSM vs. G + MSM vs. placebo	3 months	Commercial party	No contact with authors/institutions
Herrero-Beaumont *et al*[19]	Spain/Portugal	Knee OA	104	106	Crystalline GS versus placebo versus acetaminophen	6 months	Commercial party	Data sharing not allowed by study sponsor
Kwoh *et al*[22]	USA	Knee OA	103	98	GH versus placebo	6 months	Commercial party	Data sharing not allowed by study sponsor
Pavelká *et al*[25]	Czech Republic	Knee OA	101	101	Crystalline GS versus placebo	36 months	Commercial party	Data sharing not allowed by study sponsor
Reginster *et al*[27]	Belgium	Knee OA	106	106	GS versus placebo	36 months	Commercial party	Data sharing not allowed by study sponsor

*Trial identified in trial registry, no publication available.
**Long-term follow-up of Clegg *et al*.
***The first 163 patients were randomised over placebo and glucosamine hydrochloride; the remaining subjects over placebo and glucosamine sulfate.
CS, chondroitin sulfate; G, unknown which glucosamine substance; GH, glucosamine hydrochloride; GS, glucosamine sulfate; N, number of patients randomised to the specific group; MSM, methylsulfonylmethane.

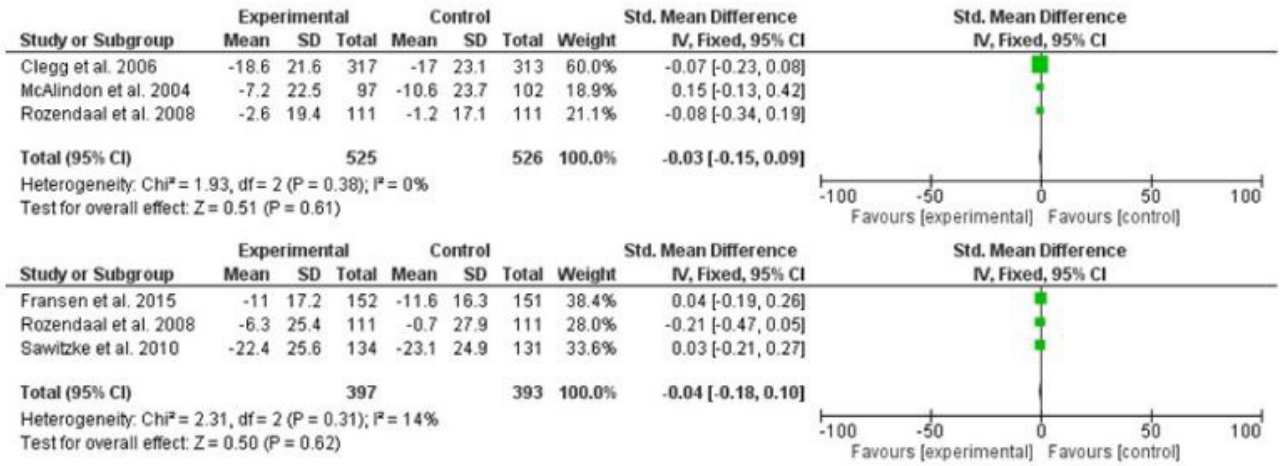


Figure 2 Forest plots for mean change in WOMAC pain at short-term (upper panel) and long-term (lower panel) on a 0–100 scale for studies that shared trial data.

long-term outcomes[16 29] and the one hip OA trial[28] with short-term and long-term outcomes. Given this lack of consistency, stratification of the subgroup analysis was done for knee OA trials only on long-term outcomes.

Risk of bias and heterogeneity

All five studies were defined as having a low risk of bias (table 2) and heterogeneity was low (I^2=0 for main effects on pain at short and I^2=14 for long-term follow-up, see figure 2), so no sensitivity analyses were performed.

Overall intervention effects

Estimated pooled differences for the primary and secondary outcome measures are presented in table 3. No statistical significance main effects were found for glucosamine over placebo.

Subgroup effects

None of the interaction terms of the predefined subgroups reached statistical significance (see table 3). Estimated pooled differences within each subgroup for the primary outcomes over all eligible trials are presented in figure 3. Within the stratified analyses among studies using GS for knee OA, the number of subjects with high baseline pain was too small for the software to test the pooled interaction term for the baseline pain severity subgroup.

DISCUSSION

To our knowledge, this is the first IPD meta-analysis to examine potential subgroup effects of oral glucosamine for people with OA. Within the five trials where the authors were willing to share their data, 1625 patients with knee or hip OA were analysed. This represents 55% of all available participants from the placebo controlled trials for this product. The main findings are: (1) overall, glucosamine was no better than placebo for both pain and function outcomes; (2) in subgroup analyses, glucosamine was no better than placebo according to baseline pain severity, BMI, gender, structural abnormalities and presence of inflammation; and (3) the majority of trials were knee OA (four trials, 1403 patients), and the analysis based on knee OA only had similar results.

Several systematic reviews and network meta-analyses have shown that as the number of high-quality and industry-independent studies on the effectiveness of glucosamine for OA increased over time, the results of earlier studies that showed beneficial effects of glucosamine were viewed as less credible.[3–5 8 34 35] It is therefore not surprising that the present IPD meta-analysis also showed no significant main effects, especially since previous studies showed a low risk of bias to be associated with small, non-significant effect sizes for glucosamine over placebo[3 8 34] and the fact that all included studies had a low risk of bias. Present results of overall treatment effects within the trials that shared data and over the different stratifications ranged from −0.43 to 2.02 on the 0–100 WOMAC pain scale, which is comparable

Table 2 Risk of bias assessment of studies included in glucosamine versus placebo comparison

	A1	B2	C3	C4	C5	D6	D7	E8	E9	E10	E11	E12	Total
Clegg *et al*[14]	+	+	+	+	+	+	+	+	+	+	+	+	Low risk
Fransen *et al*[16]	+	+	+	+	+	+	+	+	+	?	+	+	Low risk
McAlindon *et al*[24]	?	+	+	+	+	+	+	-	+	+	+	+	Low risk
Rozendaal *et al**[28]	+	+	+	+	+	+	+	+	+	+	+	+	Low risk
Sawitzke *et al*[29]	+	+	+	+	+	-	+	+	+	+	+	+	Low risk

+All values(low risk of bias); -, no (high risk of bias); ?, unclear.
A1, method of randomisation adequate; B2, treatment allocation concealed; C3, patient blinded to the intervention; C4, care provider blinded to the intervention; C5, outcome assessor blinded to the intervention; D6, drop-out rate described and acceptable; D7, randomised participants analysed in the group to which they were allocated; E8, groups similar at baseline regarding the most important prognostic indicators; E9, cointerventions avoided or similar; E10, compliance acceptable; E11, timing of the outcome assessment similar in all groups; E12, selective outcome reporting. Overall, low risk of bias was defined as fulfilling six or more of the criteria items.
*Scored by JR and MvM due to study involvement of RR.

Table 3 Estimated pooled differences (95% CI) between glucosamine and placebo on a 0–100 scale (positive values indicate a greater reduction in the outcome measure for glucosamine) and p values for treatment–subgroup interactions

		All studies (n=1625 in five studies)	Knee OA only (n=1403 in four studies)	GH in knee OA (n=1058 in three studies)	GS in knee and hip OA (n=567 in three studies)	GS in knee OA (n=345 in two studies)
Pain at short-term*	*Estimated pooled differences and 95% CI*					
	Glucosamine vs placebo	0.60 (−1.80 to 3.00)	0.91 (−1.91 to 3.75)	0.98 (−1.94 to 3.91)	−0.43 (−4.44 to 3.58)	0.59 (−11.79 to 12.98)
	p Values for treatment–subgroup interactions					
	Pain subgroup†	0.77	0.97	0.80	0.17	–‡
	BMI subgroup§	0.31	0.62	0.56	0.41	0.89
	Sex subgroup¶	0.68	0.59	0.68	0.86	0.68
	KL subgroup**	0.75	–	–	–	–
	Inflammation subgroup††	0.92	–	–	–	–
Pain at long-term*	*Estimated pooled differences and 95% CI*					
	Glucosamine versus placebo	0.98 (−1.76 to 3.73)	0.19 (−2.83 to 3.22)	0.78 (−4.33 to 5.89)	1.22 (−1.90 to 4.33)	−0.38 (−3.67 to 2.90)
	p Values for treatment–subgroup interactions					
	Pain subgroup†	0.26	0.28	0.42	0.44	0.86
	BMI subgroup‡	0.55	0.10	0.51	0.72	0.10
	Sex subgroup§	0.46	0.53	0.75	0.52	0.77
	KL subgroup¶	0.72	0.40	–	–	–
	Inflammation subgroup**	0.23	–	–	–	–
Function at short-term‡‡	*Estimated pooled differences and 95% CI*					
	Glucosamine versus placebo	1.74 (−0.45 to 3.96)	1.80 (−0.81 to 4.04)	1.92 (−0.77 to 4.61)	1.23 (−2.11 to 4.57)	−0.39 (−10.88 to 10.09)
	p Values for treatment–subgroup interactions					
	Pain subgroup†	0.47	0.34	0.37	0.69	-‡
	BMI subgroup‡	0.87	0.83	0.64	0.38	0.12
	Sex subgroup§	0.47	0.30	0.39	0.91	0.34
	KL subgroup¶	0.96	–	–	–	–
	Inflammation subgroup**	0.37	–	–	–	–
Function at long-term‡‡	Estimated pooled differences and 95% CI					
	Glucosamine versus placebo	1.40 (−1.27 to 4.06)	0.63 (−2.31 to 3.58)	0.85 (−4.43 to 6.13)	2.02 (−0.82 to 4.86)	0.62 (−2.29 to 3.52)
	p Values for treatment-subgroup interactions					
	Pain subgroup†	0.49	0.38	0.55	0.94	0.91
	BMI subgroup‡	0.82	0.42	0.65	0.56	0.68
	Sex subgroup§	0.72	0.61	0.80	1.00	0.94
	KL subgroup¶	0.83	0.77	–	–	–
	Inflammation subgroup**	0.46	–	–	–	–

*Measured using WOMAC pain (0–100) and adjusted for age sex, BMI, WOMAC pain at baseline and study number.
†WOMAC pain <70 versus ≥70 on a 0–100 scale.
‡Too few cases in high pain group for the software to test the interaction term.
§BMI <27 kg/m^2 versus ≥27 kg/m^2.
¶Male versus female.
**Kellgren and Lawrence grades 0–2 versus 3–4 (not available in McAlindon *et al*).[24]
††Measured using WOMAC function (0–100) and adjusted for age sex, BMI, WOMAC function at baseline and study number. Positive estimated pooled differences indicate a greater reduction in the outcome in the glucosamine group compared with the placebo group.
‡‡Presence of inflammation, defined as presence of swelling/effusion on clinical examination or an elevated ESR, defined as ESR ≥20 mm/h for men aged ≥50 years, ESR ≥15 mm/h for men aged <50 years, ESR ≥30 mm/h for women aged ≥50 years and ESR ≥20 mm/h for women aged <50 years, versus absence of inflammation (not available in McAlindon *et al*[24] and Fransen *et al*).[16]
BMI, body mass index; ESR, erythrocyte sedimentation rate; GH, glucosamine hydrochloride; GS, glucosamine sulfate; KL, Kellgren and Lawrence grade; OA, osteoarthritis.

with the overall treatment effects for industry independent studies (0.1 (95% CI −0.2 to 0.5)) for visual analog scale (VAS) pain on a 0–10 scale) presented by the meta-analyses of Wandel and colleagues.[3] In the literature, overall beneficial effects of treatment have been reported in studies using the glucosamine compound produced by Rottapharm[4 5 8 34]; however, these trials were not made available to the study team for the current analyses.

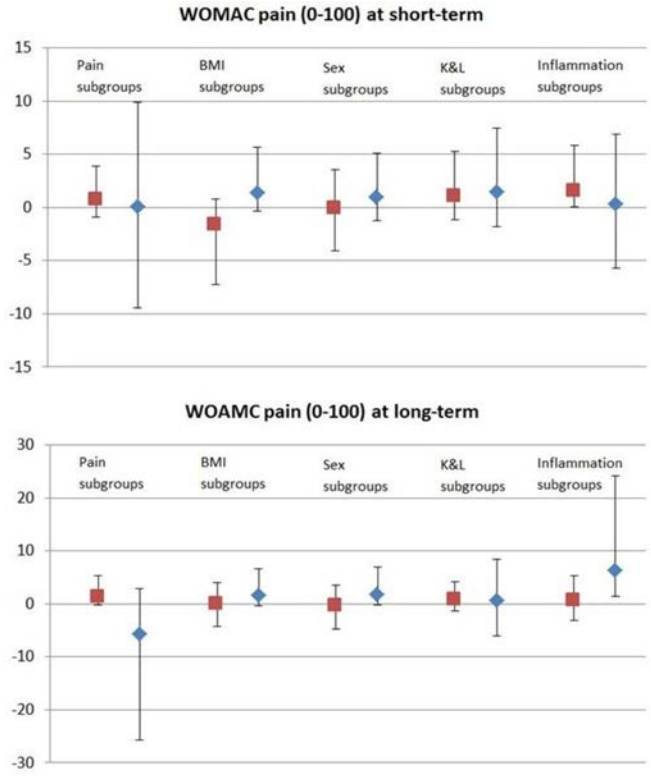


Figure 3 Estimated pooled differences between glucosamine and placebo within predefined subgroups for all eligible trials. Positive values indicate a greater reduction in the outcome measure for glucosamine. Red figures represent low pain (WOMAC pain <70), low BMI (<27 kg/m^2), male sex, K&L grades 0–2 and absence of inflammation subgroups, respectively. Blue figures represent high pain (WOMAC pain ≥70), high BMI (≥27 kg/m^2), female sex, K&L grades 3–4 and presence of inflammation subgroups, respectively. BMI, body mass index; K&L, Kellgren and Lawrence grade.

Extending previous initiatives, the present study also evaluated treatment effects of glucosamine over placebo for several clinically relevant subgroups of OA, made possible by the IPD from the collaborating trials. Despite the large number of participants incorporated in the IPD meta-analysis, none of the interaction terms reached statistical significance. The interactions with BMI among knee OA patients receiving GS on short-term function (p=0.12) and on long-term pain (p=0.10) were the only outcomes for which further research may be warranted. However, given the number of analyses performed in the study, incidental findings are certainly possible.

The currently used cut-off for the baseline pain severity subgroup is somewhat comparable with the strata used in the Clegg *et al* study[14] to test for different effects within subjects with mild pain (WOMAC pain scores 0–60) versus those with moderate to severe pain (WOMAC pain scores 60–80). The Clegg *et al* study was not powered to show subgroup effects, but the non-significant effects of glucosamine over placebo within both subgroups is corroborated by the present results.

The current study has several limitations. Despite all efforts, data from only six of the 21 identified studies were acquired. Of those studies not included in the present study, the largest groups were those not responding to any of the requests for data sharing (six studies) and those not permitted by the commercial study sponsor to share data (four studies) (see table 1). Although missing data for the main baseline characteristics within the data shared with the OA Trial Bank were limited, multiple imputation methods were needed to deal with the missing data in the outcome measures that ranged from 2% to 46%. Within the trials that shared data, only a few measured the predefined subgroups based on structural abnormalities and presence of inflammation. The available data for these subgroups combined studies evaluating different glucosamine substances for different OA joints. Therefore, rigorous stratification of the analysis was not possible with the available data.

Open access to data of clinical trials has been proposed and debated for several years.[36-38] Nevertheless our experience, in common with others, suggests that currently this is far from accepted practice.[39] Thus, the full potential and use of completed clinical trials is not reached and only part of the clinical evidence is available to clinicians and patients, thus threatening the appropriateness of recommendations for clinical decision making.[39] Once initiatives such as the OA Trial Bank, which appropriately use existing data for scientific purposes, become more established and generally accepted, authors and commercial parties involved in clinical research may become more confident in data sharing. The OA Trial Bank plans to update publications every 5 years and will again approach data owners that chose to not share their data to the OA Trial Bank in the first initiative.

The aim of the present study was to perform an IPD meta-analysis on all available RCTs on glucosamine in people with OA. After performing the systematic search of the literature and clinical trial registers, it took 18 months to reach as many data owners as possible and to collect and check all data of those willing to deliver their trial data. For a systematic review, one might argue that an update of the search strategy is warranted. However, given the time-consuming efforts of sharing data between research institutes, this was not feasible for the present study.

In conclusion, the current IPD on the efficacy of glucosamine for subgroups of OA based on pain severity, BMI, sex, radiographic structural changes and presence of inflammation, using data from 55% of the participants available in literature and using data from low risk-of-bias trials only, did not identify a subgroup for which glucosamine showed any significant beneficial effects over placebo for pain or function in either the short term or long term. Stratification only for participants with knee OA or for type of glucosamine did not result in any differences in outcomes. Therefore, currently, there is no evidence to support the use of glucosamine for treatment of hip or knee OA in general and an absence of evidence to support the use of glucosamine for clinically relevant subgroups of OA according to baseline pain severity, BMI, sex, structural abnormalities and presence of inflammation.

Author affiliations
[1]Department of General Practice, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands
[2]Department of Rheumatology & Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands
[3]Department of Rheumatology, The University of Nottingham, Nottingham, UK
[4]Department of Primary Care & Health Services, Keele University, Staffordshire, UK
[5]Department of Clinical Sciences, Orthopaedics, Lund University, Lund, Sweden
[6]Department of Rheumatology, Tufts Medical Center, Boston, Massachusetts, USA
[7]Department of General Practice and Orthopaedics, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands

Acknowledgements All data owners (researchers and institutes) are to be acknowledged for providing their data to the OA Trial Bank.

Contributors JR, MvM, HJWB, MD, KSD, LSL, TM, WZ and SBZ have substantially contributed to the conception and design of the work. JR and RMR have substantially contributed to the acquisition of the data and drafted the manuscript. All authors contributed to the analysis and interpretation of the work, revised the manuscript critically for important intellectual content, approved the

final version of the manuscript and agreed to be accountable for all aspects of the work.

Funding The current project received funding by the Dutch Arthritis Foundation (BP12-1-161). KSD is part funded by the National Institute for Health Research (NIHR) Collaborations for Leadership inApplied Research and Care West Midlands and by a Knowledge Mobilisation Research Fellowship (KMRF-2014-03-002) from the NIHR. This paper presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. WZ is supported by a grant from Arthritis Research UK. TEM is supported by grants from Sanofi Aventis, Abbvie, Fidia, Samumed and Pfizer and personal fees from Flexion Therapeutics, Samumed, Plexxikon Inc, Regeneron, Orthogen and McNeil Consumer HC. LSL is supported by personal fees from Galapagos NV, FlexionTherapeutics, Johnson & Johnson, Regeneron, Össur and Samumed.

Competing interests None declared.

Patient consent Two representatives of patient and public involvement (members of the Arthritis research UK OA Research Users Group) are official members of the Steering Committee of the OA Trial Bank. These representatives provided feedback on the design of the study, including study selection, selection and definitions of subgroups, and outcome measures. Also for dissemination activities of OA Trial Bank and for prioritisation of future research questions, the input from patient and public involvement is obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Zhang W, Nuki G, Moskowitz RW, *et al*. OARSI recommendations for the management of hip and knee osteoarthritis: part III: changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis Cartilage* 2010;18:476–99.
2 National Clinical Guideline C. 2014.
3 Wandel S, Jüni P, Tendal B, *et al*. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *BMJ* 2010;341:c4675.
4 Vlad SC, LaValley MP, McAlindon TE, *et al*. Glucosamine for pain in osteoarthritis: why do trial results differ? *Arthritis Rheum* 2007;56:2267–77.
5 McAlindon TE, LaValley MP, Gulin JP, *et al*. Glucosamine and chondroitin for treatment of osteoarthritis: a systematic quality assessment and meta-analysis. *JAMA* 2000;283:1469–75.
6 Bierma-Zeinstra SM, Verhagen AP. Osteoarthritis subpopulations and implications for clinical trial design. *Arthritis Res Ther* 2011;13:213.
7 Zhang W, Doherty M, Arden N, *et al*.EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR standing Committee for International clinical studies including therapeutics (ESCISIT). *Ann Rheum Dis* 2005;64:669–81.
8 Towheed TE, Maxwell L, Anastassiades TP, *et al*. Glucosamine therapy for treating osteoarthritis. *Cochrane Database Syst Rev* 2005:CD002946. CD002946..
9 van Middelkoop M, Dziedzic KS, Doherty M, *et al*. Individual patient data meta-analysis of trials investigating the effectiveness of intra-articular glucocorticoid injections in patients with knee or hip osteoarthritis: an OA Trial Bank protocol for a systematic review. *Syst Rev* 2013;2:54.
10 van Middelkoop M, Arden NK, Atchia I, *et al*. The OA Trial Bank: meta-analysis of individual patient data from knee and hip osteoarthritis trials show that patients with severe pain exhibit greater benefit from intra-articular glucocorticoids. *Osteoarthritis Cartilage* 2016;24:1143–52.
11 Altman R, Asch E, Bloch D, *et al*. Development of criteria for the classification and reporting of osteoarthritis. classification of osteoarthritis of the knee. diagnostic and therapeutic Criteria Committee of the american Rheumatism Association. *Arthritis Rheum* 1986;29:1039–49.
12 Chopra A, Saluja M, Tillu G, *et al*. A Randomized Controlled Exploratory evaluation of standardized ayurvedic formulations in symptomatic osteoarthritis knees: A government of India NMITLI Project. *Evid Based Complement Alternat Med* 2011;2011:1–12.
13 Cibere J, Kopec JA, Thorne A, *et al*. Randomized, double-blind, placebo-controlled glucosamine discontinuation trial in knee osteoarthritis. *Arthritis Rheum* 2004;51:738–45.
14 Clegg DO, Reda DJ, Harris CL, *et al*. Glucosamine, chondroitin sulfate, and the two in combination for painful knee osteoarthritis. *N Engl J Med* 2006;354:795–808.
15 Coulson S, Butt H, Vecchio P, *et al*. Green-lipped mussel extract (Perna canaliculus) and glucosamine sulphate in patients with knee osteoarthritis: therapeutic efficacy and effects on gastrointestinal Microbiota profiles. *Inflammopharmacology* 2013;21:79–90.
16 Fransen M, Agaliotis M, Nairn L, *et al*. Glucosamine and chondroitin for knee osteoarthritis: a double-blind randomised placebo-controlled clinical trial evaluating single and combination regimens. *Ann Rheum Dis* 2015;74:851–8.
17 Frestedt JL, Walsh M, Kuskowski MA, *et al*. A natural mineral supplement provides relief from knee osteoarthritis symptoms: a randomized controlled pilot trial. *Nutr J* 2008;7:9.
18 Hatano K, Miyakuni Y. Effects and safety of Soymilk Beverage Containing N-acetyl glucosamine on osteoarthritis. *Japan Pharmacology & Therapeutics* 2006;34:149–65.
19 Herrero-Beaumont G, Ivorra JA, Del Carmen Trabado M, *et al*. Glucosamine sulfate in the treatment of knee osteoarthritis symptoms: a randomized, double-blind, placebo-controlled study using acetaminophen as a side comparator. *Arthritis Rheum* 2007;56:555–67.
20 Hughes R, Carr A. A randomized, double-blind, placebo-controlled trial of glucosamine sulphate as an analgesic in osteoarthritis of the knee. *Rheumatology* 2002;41:279–84.
21 Kawasaki T, Kurosawa H, Ikeda H, *et al*. Additive effects of glucosamine or risedronate for the treatment of osteoarthritis of the knee combined with home exercise: a prospective randomized 18-month trial. *J Bone Miner Metab* 2008;26:279–87.
22 Kwoh CK, Roemer FW, Hannon MJ, *et al*. Effect of oral glucosamine on joint structure in individuals with chronic knee pain: a randomized, placebo-controlled clinical trial. *Arthritis Rheumatol* 2014;66:930–9.
23 Martí-Bonmatí L, Sanz-Requena R, Rodrigo JL, *et al*. Glucosamine sulfate effect on the degenerated patellar cartilage: preliminary findings by pharmacokinetic magnetic resonance modeling. *Eur Radiol* 2009;19:1512–8.
24 McAlindon T, Formica M, LaValley M, *et al*. Effectiveness of glucosamine for symptoms of knee osteoarthritis: results from an internet-based randomized double-blind controlled trial. *Am J Med* 2004;117:643–9.
25 Pavelká K, Gatterová J, Olejarová M, *et al*. Glucosamine sulfate use and delay of progression of knee osteoarthritis: a 3-year, randomized, placebo-controlled, double-blind study. *Arch Intern Med* 2002;162:2113–23.
26 Petersen SG, Beyer N, Hansen M, *et al*. Nonsteroidal anti-inflammatory drug or glucosamine reduced pain and improved muscle strength with resistance training in a randomized controlled trial of knee osteoarthritis patients. *Arch Phys Med Rehabil* 2011;92:1185–93.
27 Reginster JY, Deroisy R, Rovati LC, *et al*. Long-term effects of glucosamine sulphate on osteoarthritis progression: a randomised, placebo-controlled clinical trial. *Lancet* 2001;357:251–6.
28 Rozendaal RM, Koes BW, van Osch GJ, *et al*. Effect of glucosamine sulfate on hip osteoarthritis: a randomized trial. *Ann Intern Med* 2008;148:268–77.
29 Sawitzke AD, Shi H, Finco MF, *et al*. Clinical efficacy and safety of glucosamine, chondroitin sulphate, their combination, celecoxib or placebo taken to treat osteoarthritis of the knee: 2-year results from GAIT. *Ann Rheum Dis* 2010;69:1459–64.
30 Usha PR, Naidu MU. Randomised, Double-Blind, Parallel, Placebo-Controlled Study of Oral Glucosamine, Methylsulfonylmethane and their combination in osteoarthritis. *Clin Drug Investig* 2004;24:353–63.
31 Giordano N, Fioravanti A, Papakostas P, *et al*. The efficacy and tolerability of glucosamine sulfate in the treatment of knee osteoarthritis: A randomized, double-blind, placebo-controlled trial. *Curr Ther Res Clin Exp* 2009;70:185–96.
32 Bellamy N, Buchanan WW, Goldsmith CH, *et al*. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833–40.
33 Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;16:494–502.
34 Eriksen P, Bartels EM, Altman RD, *et al*. Risk of Bias and brand explain the observed inconsistency in trials on glucosamine for symptomatic relief of osteoarthritis: a meta-analysis of placebo-controlled trials. *Arthritis Care Res* 2014;66:1844–55.
35 Nüesch E, Trelle S, Reichenbach S, *et al*. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 2009;339:b3244.
36 Lehman R, Loder E. Missing clinical trial data. *BMJ* 2012;344:d8158.
37 Loder E. Sharing data from clinical trials: where we are and what lies ahead. *BMJ* 2013;347:f4794.
38 Koenig F, Slattery J, Groves T, *et al*. Sharing clinical trial data on patient level: opportunities and challenges. *Biom J* 2015;57:8–26.
39 Fleetcroft R, Ford J, Gollop ND, *et al*. Difficulty accessing data from randomised trials of drugs for heart failure: a call for action. *BMJ* 2015;351:h5002.

OPEN ACCESS

EXTENDED REPORT

Weight loss for overweight and obese individuals with gout: a systematic review of longitudinal studies

Sabrina M Nielsen,[1] Else M Bartels,[1] Marius Henriksen,[1,2] Eva E Wæhrens,[1,3] Henrik Gudbergsen,[1] Henning Bliddal,[1] Arne Astrup,[4] Filip K Knop,[5,6,7] Loreto Carmona,[8] William J Taylor,[9] Jasvinder A Singh,[10] Fernando Perez-Ruiz,[11] Lars E Kristensen,[1] Robin Christensen[1]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211472).

For numbered affiliations see end of article.

Correspondence to
Robin Christensen, Biostatistician Professor of clinical epidemiology, adj Head of Musculoskeletal Statistics Unit, The Parker Institute,Bispebjerg and Frederiksberg Hospital, Nordre Fasanvej 57DK-2000 Copenhagen Frederiksberg, Denmark;
robin.christensen@regionh.dk

Received 15 March 2017
Revised 29 May 2017
Accepted 1 July 2017
Published Online First
2 September 2017

ABSTRACT

Objectives Weight loss is commonly recommended for gout, but the magnitude of the effect has not been evaluated in a systematic review. The aim of this systematic review was to determine benefits and harms associated with weight loss in overweight and obese patients with gout.

Methods We searched six databases for longitudinal studies, reporting the effect of weight loss in overweight/obese gout patients. Risk of bias was assessed using the tool Risk of Bias in Non-Randomised Studies of Interventions. The quality of evidence was assessed using the Grading of Recommendations Assessment, Development and Evaluation.

Results From 3991 potentially eligible studies, 10 were included (including one randomised trial). Interventions included diet with/without physical activity, bariatric surgery, diuretics, metformin or no intervention. Mean weight losses ranged from 3 kg to 34 kg. Clinical heterogeneity in study characteristics precluded meta-analysis. The effect on serum uric acid (sUA) ranged from −168 to 30 µmol/L, and 0%–60% patients achieving sUA target (<360 µmol/L). Six out of eight studies (75%) showed beneficial effects on gout attacks. Two studies indicated dose–response relationship for sUA, achieving sUA target and gout attacks. At short term, temporary increased sUA and gout attacks tended to occur after bariatric surgery.

Conclusions The available evidence is in favour of weight loss for overweight/obese gout patients, with low, moderate and low quality of evidence for effects on sUA, achieving sUA target and gout attacks, respectively. At short term, unfavourable effects may occur. Since the current evidence consists of a few studies (mostly observational) of low methodological quality, there is an urgent need to initiate rigorous prospective studies (preferably randomised controlled trials).

Systematic review registration PROSPERO, CRD42016037937.

INTRODUCTION

Gout is a common form of inflammatory arthritis,[1 2] with an age-standardised global prevalence of 0.08% and is higher in developed countries.[3] Gout is a crystal-deposition disease resulting from chronic elevation of serum uric acid (sUA) above the saturation point for monosodium urate (MSU).[4–7] Initial presentation is severely painful episodes of peripheral joint synovitis (acute 'attacks'), but joint damage and subcutaneous tophus deposition may develop.[8] The general management principle is to reduce sUA levels, allowing MSU crystals to dissolve, leading to the elimination of acute attacks, disappearance of tophi and possibly cure of the disease.[9–11]

Body mass index (BMI) is strongly positively correlated to sUA levels,[12 13] and weight loss is a commonly recommended treatment for gout.[14–23] Furthermore, weight loss from bariatric surgery is associated with reduced incidence of hyperuricaemia and gout.[24] The mechanism by which weight loss can lower sUA levels is poorly understood. Some suggest that improved insulin resistance results in less insulin-enhanced reabsorption of organic anions such as urate,[2] and a study demonstrated decreased sUA in overweight patients receiving either weight loss from low-energy diet or an insulin-sensitising agent.[25] However, a study of severe obese patients receiving bariatric surgery found no association between reduced sUA levels and improved insulin resistance,[26] making a relationship questionable.

Guidelines recommending weight loss for gout patients[14–23] are based on evidence from only few clinical studies,[27 28] one population-based study[29] and indirect evidence from studies on non-gout subjects. The evidence for effectiveness in clinical studies has to our knowledge not previously been evaluated in a systematic review. Therefore, the primary objective of this systematic review was to determine the benefits and harms associated with weight loss in overweight and obese individuals with gout. Furthermore, we had an explicit focus on the weight loss intervention (including magnitude and intensity) to see whether a dose–response relationship exists at the study (ie, group) level.

METHODS

Protocol

A protocol adhering to the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols 2015 statement[30] was registered online (PROSPERO: CRD42016037937) and published on www.parkerinst.dk.

Search strategy

We searched four bibliographic databases on 26 April 2016; MEDLINE via Ovid from 1946, EMBASE via Ovid from 1974, Web of Science via Web of Knowledge from 1900, Cochrane Central

To cite: Nielsen SM, Bartels EM, Henriksen M, *et al*. *Ann Rheum Dis* 2017;**76**:1870–1882.

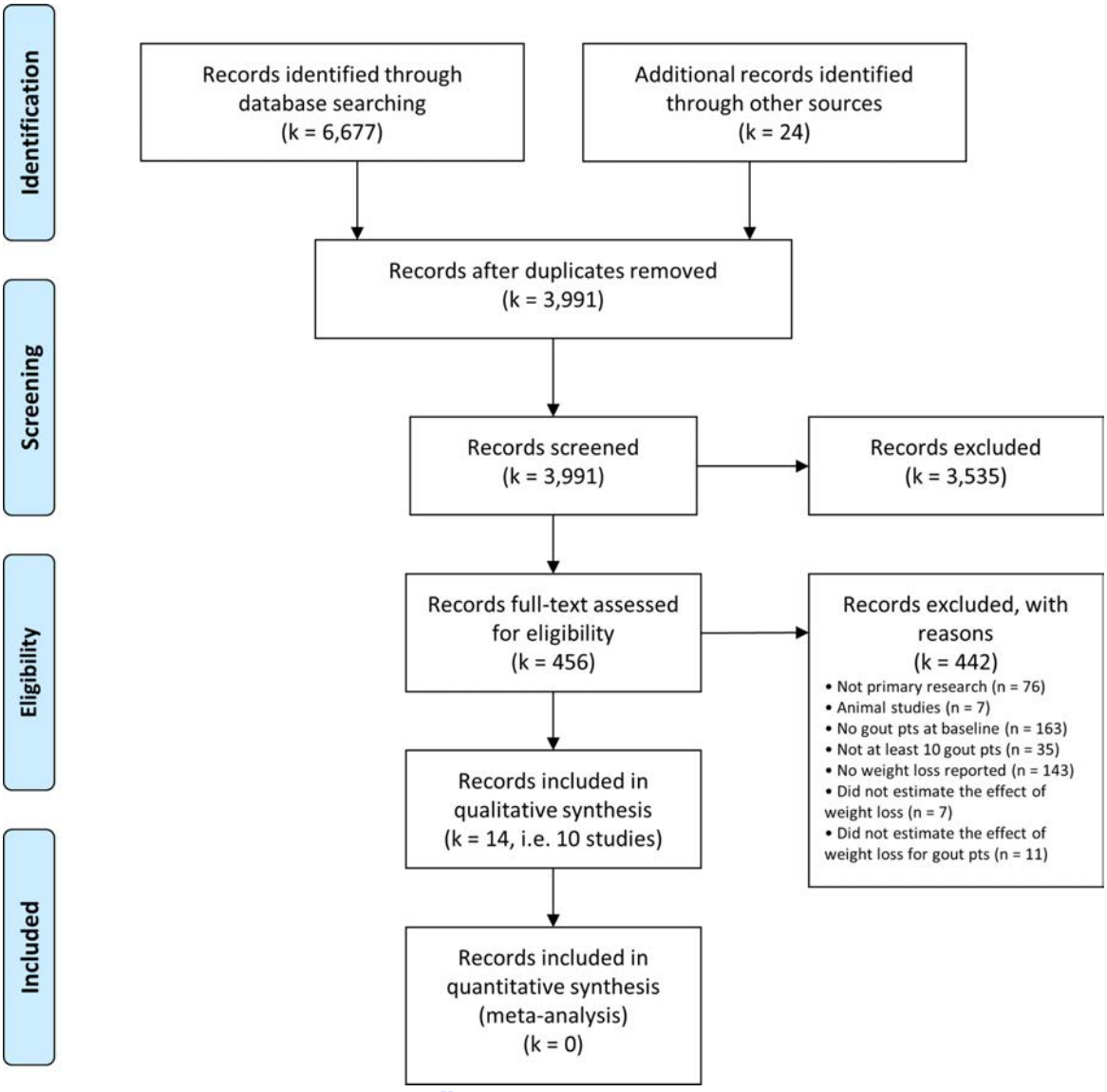


Figure 1 PRISMA flow diagram. Modified from Moher *et al*.[75] pts, patients.

Register of Controlled Trials (CENTRAL, The Cochrane Library), as well as ClinicalTrials.gov and WHO International Clinical Trial Registry Platform portal (search strategy presented: online supplementary text S1). We screened reference lists of relevant articles, as well as the American College of Rheumatology (ACR) and European League Against Rheumatism conference abstracts from 2014 and 2015, the ACR conference abstracts from 2016 and content experts were asked if they were aware of any other relevant studies.

Study selection

Anticipating only few randomised controlled trials (RCTs), we also included longitudinal observational studies (non-randomised studies) that quantitatively estimated the effect following weight loss. Studies needed to include ≥10 adult and overweight/obese patients (author described or BMI ≥25 kg/m^2[31]) with diagnosed gout (author described or meeting the 1977 ACR criteria for gout[32]). Eligible interventions included those where a weight reduction was reported explicitly, whether intentional or unintentional. The weight reduction was required to be the only difference in terms of intervention from the defined control group. Two reviewers (SMN supported by EMB) assessed the records. Disagreements were resolved by consensus or by discussion with a third reviewer (RC).

Data extraction and management

Two reviewers (SMN supported by RC) extracted the data. Prespecified outcomes included essential outcome domains for chronic gout[33 34]: (1) joint pain; (2) tophus/tophi; (3) physical function; (4) health-related quality of life (HRQoL); (5) sUA change; (6) Achieving sUA target (ie, sUA reduction to <360 μmol/L (6 mg/dL)); (7) serious adverse events (SAEs, defined as adverse events that are fatal, life-threatening or require hospitalisation); (8) withdrawals due to adverse events (WDdtAEs); (9) patient global assessment; (10) wody weight change; and (11) gout attacks (any measure).

Assessment of risk of bias in included studies (internal validity)

Two reviewers (SMN supported by RC) assessed risk of bias using the tool Risk of Bias In Non-Randomised Studies of Interventions[35 36] for evaluation of the risk of bias in non-randomised studies comparing health effects of two or more interventions.

Table 1 Characteristics of the included studies

	Study characteristics		Characteristics of gout pts						
Author, year, (multiple publication)	Study population	Intervention	No. of pts	Females, n pts	Age, years	BMI, kg/m^2	Disease duration, years	Urate-lowering medication use, n pts	Presence of tophi, n pts
Nguyen *et al*, 2016 [44] (Sherwin *et al* 1981 [47])	Men with a high cardiovascular risk profile and no gout at baseline of the MRFIT* study (n=11 896). 408 developed gout during the 7-year follow-up period.	This substudy of MRFIT uses a subpopulation, stratifying the risk of recurrent gout attacks for the BMI change in gout pts.	408†	0 (0)	NA	NA	NA	NA	NA
Dalbeth *et al*, 2014 [27] (Dalbeth, *et al* 2013 [42])	Pts with T2D and BMI ≥35 kg/m^2 who met presurgery weight loss requirements during part 1 (72 included, 60 completed). Included gout pts.	(Part 1) Presurgery weight loss (dietetic advice, set goals to establish a regular exercise programme and lose 5–10 kg, in addition to Optifast VLCD programme 4 weeks prior to surgery). (Part 2) Bariatric surgery (laparoscopic sleeve gastrectomy)	12	5 (42)	49 (8)	48.5 (5.4)	6.4 (7)	9 (75)	NA
Romero-Talamás *et al*, 2014 [46]	Gout pts with obesity (n=155).	I: Bariatric surgery (Roux-en-Y gastric bypass, gastrectomy, or adjustable gastric banding)	99	75 (75)	52.1 (10.3)	49.5 (11.9)	NA	53 (56)	NA
		C: Non-bariatric surgery (laparoscopic cholecystectomy, open cholecystectomy, cholecystectomy with another contaminant procedure or laparoscopic Heller myotomy).	56	34 (61)	63.3 (11.9)	36.8 (9.8)	NA	23 (42)	NA
Zeng *et al*, 2012 [49] *(In Chinese, RCT)*	Overweight men with gout, not using urate-lowering medication or having tophi (67 randomised, 61 completed the study).	I: High protein	30	0 (0)	61.5 (14.5)	27.0 (1.3)	NA	0 (0)	0 (0)
		C: Low purine	31	0 (0)	61.3 (11.2)	27.2 (1.6)	NA	0 (0)	0 (0)
Perez-Ruiz *et al*, 2011 [45] (*unpublished data*)	Gout pts with 5 years of compliance to urate-lowering therapy, and no tophi or resolution of all tophi prior to withdrawing urate-lowering therapy (n=211).	Withdrawal of urate-lowering therapy, stratified according to weight loss: lost >5% weight	25‡	1 (4)	NA	NA	7.4 (6.4)	25 (100)	0 (0)
		Withdrawal of urate-lowering therapy, stratified according to weight loss: no weight loss	167‡	7 (4)	NA	NA	7.2 (6.4)	167 (100)	0 (0)
Zhu *et al*, 2010 [29] (Sherwin, *et al* 1981 [47])	Men with a high cardiovascular risk profile and sufficient data from the MRFIT* study (n=12 379). Included gout pts.	This substudy of MRFIT uses a subpopulation, stratifying the changes in sUA for baseline BMI, weight change and other variables.	NA†	0 (0)	NA (All: 46 (range, 35–57))	NA	NA	NA	NA
Barskova *et al*, 2009 [40] *(In Russian)*	Gout pts (30 included, 23 completed).	Metformin (1500 mg/day)	23	2 (7)	51 (range, 43–54)	32.6 (5.2)	6 (range, 4–11)	0 (0)	8 (27)
Friedman *et al*, 2008 [43]	Pts received gastric bypass and experiencing postoperative gout attacks.	Bariatric surgery (gastric bypass and preoperative preparation, that is, clear liquid diet with protein supplementation 4–7 days and mechanical bowel preparation 2 days).	21	6 (29)	52 (range, 32–73)	49.6 (range, 36.1– 63)	NA	NA	NA
Dessein *et al*, 2000 [28] (Terkeltaub, *et al*, 2001 [48])	Gout pts. without tophi (n=13).	Diet recommendations (calorie restriction with specified macronutrient proportions, replacing refined carbohydrates with complex ones and replacing saturated fats with monosaturated and polyunsaturated ones).	13	0 (0)	50§ (5.6; range, 38–62)	30.5§¶ (8.1)	7§(10.2; range, 0.5–38)	0 (0)	0 (0)
Brandstetter *et al*, 1986 [41] (*In German*)	Gout pts with hyptertension (n=22)	I: Allopurinol, diet (low purine) and celiprolol (betablocker) and chlorthalidone (diuretic). C: Allopurinol, diet (low purine) and celiprolol .	11+11 (both groups)	7 (32)	52§ (range, 28–68)	26.0**	NA	NA	NA

The results are reported as mean (SD) or number (%), unless otherwise indicated.
*The original study (MRFIT) [39] randomised pts (n=12 866) to intervention (smoking cessation, weight reduction by caloric intake reduction and increased physical activity, nutritional counselling and antihypertensive treatment) or control.
†May include gout pts that were not overweight at baseline.
‡Due to loss of data in the study, the number of pts in the groups were 25 and 167 at baseline, and 29 and 163 at latestfollow-up.
§Median.
¶One had normal BMI.
**No BMI were reported, so BMI were calculated from data on height (median, 179 (range, 162–199)), and weight (median, 83.2 (range, 58–105)).
Allopurinol, diet (low purine) and caliprolol.
BMI, body mass index; C, control treatment; I, intervention; MRFIT, The Multiple Risk Factor Intervention Trial; NA, no data available; pts, patients; RCT, randomised controlled trial; sUA, serum uric acid; T2D, type 2 diabetes.

Post hoc, we decided to also use this tool for assessing studies with only one study group, by assuming that a virtual control group not receiving any intervention and experiencing no effect on any outcome was available, and for assessing RCTs, making comparisons possible. We resolved disagreements by discussion.

Important confounders of interest and cointerventions possibly affecting the effect of weight loss were not specified at protocol stage but prior to the risk of bias assessment (online supplementary text S2).

Reporting bias in individual studies was further investigated by comparing the constructed outcome reporting matrix,[37] with the protocols (if available).

Statistical analyses and evidence synthesis

None of our planned meta-analyses were conducted due to indisputable clinical heterogeneity in study characteristics (PICOTs). Instead it was decided post hoc that data for each study would be presented for all time points in a summary of findings table and the latest time point as changes from baseline would be summarised for each study in a summary of findings and GRADE evidence profile table. Based on the tables, we qualitatively considered the impact of follow-up time, that is, short-term (<3 months), medium (3–12 months) and long-term (>12 months), acute versus chronic gout, presence versus absence of concurrent urate-lowering medication use, presence versus absence of tophi, the dose–response phenomena of weight loss in magnitude and intensity (ie, magnitude over time) and the impact of bias. A graph showing the relationship between weight loss and sUA was constructed post hoc.

Dealing with missing data

Where data were missing or incomplete, we searched for information from the study authors and from additional records for the study. No imputations were carried out for patients lost at follow-up. Missing body weights were estimated from BMI, assuming a height of 1.70 m. Missing SDs were calculated from other statistics such as standard errors, or estimated from other studies investigating gout patients; for sUA, we used a SD for change from baseline of 137 μmol/L.

Assessing the quality of the evidence

We assessed the quality of the evidence with the GRADE approach,[38] starting at low quality of evidence, since the evidence was primarily based on observational studies and subsequently down-rated or up-rated the evidence.

RESULTS

Study selection

We identified 3991 records after removal of duplicates, forwarding 456 for full-text assessments after screening (figure 1). After excluding 442 records (see online supplementary text S3), we identified 14 records describing 10 studies for inclusion in the systematic review.[27–29 39–49]

During the study selection and data extraction, authors of 18 studies[27–29 40 43–46 49–58] were contacted; three responded[27 44 45] and provided additional information, including unpublished data for Perez-Ruiz *et al*.[45]

Study characteristics

The studies were comprised of one RCT[49] and nine non-randomised studies (table 1). Gout patients were a subgroup in three of the studies,[27 29 44] of which one study initially only included non-gout patients but did a subanalysis on recurrent gout attacks for those who developed gout during follow-up.[44] The studies included between 12 and 408 gout patients, including 0%–75% females. The average age and BMI ranged from 49 to 63.3 years, and 26.0 to 49.6 kg/m^2, respectively. Case definitions of gout included the use of the 1977 ACR criteria in one study,[42] diagnosis confirmed by detecting crystals in three studies,[28 40 45] asking 'Have you been told by your physician that you have gout?' in two studies,[29 44] medical history and documentation of previous gout attacks in one study,[43] documented episode(s) or evidence of medication use in one study,[46] or not specified in two studies[41 49] (online supplementary table S1). Comorbidities selected in the studies, besides overweight, included type 2 diabetes,[27] hypertension[41] and a high cardiovascular risk profile.[29 44]

Interventions included intentional weight loss from dietary changes with or without increased physical activity,[27 28] bariatric surgery[27 43 46] and unintentional weight loss from high protein diet,[49] diuretics[41] and metformin.[40] Three studies[29 44 45] stratified according to weight or BMI reduction, using no reduction as control. Four studies had no control group.[27 28 40 43] Follow-up ranged from 4 weeks to 7 years, and a mean weight loss of 3–34 kg at latest follow-up was reported.

Effect of weight loss

No data were available for joint pain, HRQoL or patient global assessment (outcome matrix: online supplementary table S2). One study[45] provided data on tophi, reporting none for both groups at baseline and follow-up, and one study[27] provided data on physical function measured by Short Form-36 physical functioning domain, reporting diminished function with the values 43.3 (SD 21.8), 24.6 (SD 28.2), 10.8 (SD 12.8) at baseline, 6 months and 1.5 years, respectively. One study[40] reported four WDdtAEs from metformin, and one study[27] did not report any SAEs in gout patients. On sUA, achieving sUA target and gout attacks, eight, five, and eight studies provided data, respectively (table 2).

The effect on mean sUA ranged from −168 μmol/L to 30 μmol/L (−2.8 mg/dL to 0.5 mg/dL) at latest follow-up (table 3). Studies with the largest (and fastest) weight loss showed in general the largest decrease (figure 2).[27 28 46] Furthermore, a dose–response relationship was shown by Zhu *et al*[29] with a weight loss of ≥10 kg being associated with a change in sUA of −37 μmol/L. It should be noted that non-gout and non-overweight patients were included in their analysis as well. At short term, Dalbeth *et al* (part 2)[27] reported an immediate postoperative mean sUA of 510 (SD 130) μmol/L, that is, an increase of 70 μmol/L from bariatric surgery, and at latest follow-up, sUA had dropped to 330 (SD 90) μmol/L. In that period, three out of seven patients terminated urate-lowering medication, that is, the decrease may truly be larger. Three studies showed no effect on sUA; Perez-Ruiz *et al*[45] and Dalbeth *et al* (part 1),[27] both with a concurrent decrease in urate-lowering medication, and Brandstetter *et al*,[41] where the weight loss may partly be due to diuretics and hence truly lower. Barskova *et al*[40] showed a decrease in sUA from a weight loss of only 3 kg. However, the use of metformin can have affected the results.

The proportion achieving sUA target (<360 μmol/L) ranged from 0% to 60% reduction in patients with raised sUA. Furthermore, a dose–response relationship was shown by Zhu *et al*[29] with approximately three times higher odds of achieving sUA target with loss of ≥10 kg body weight during 7 years compared with not losing weight. It should be noted that non-overweight gout patients were included in their analysis as well. The 0% and

Table 2 Study findings

Author, year, (multiple publication)	Group: time point	Body weight, kg	Body weight change from baseline, kg	sUA, µmol/L	Achieving sUA target*, n pts	Gout attacks
Nguyen *et al*, 2016[44] (Sherwin *et al* 1981[47])	<−5% BMI: 12 months	NA	NA	NA	NA	Recurrent, OR 0.61 (0.32 to 1.16)†
	−3.6 to −5% BMI: 12 months	NA	NA	NA	NA	Recurrent, OR 0.94 (0.43 to 2.06)†
	No change: 12 months	NA	NA	NA	NA	Recurrent, OR 1.00 (reference)†
	+3.6 to +5% BMI: 12 months	NA	NA	NA	NA	Recurrent, OR 1.43 (0.75 to 2.72)†
	>+5% BMI: 12 months	NA	NA	NA	NA	Recurrent, OR 1.60 (0.89 to 2.89)†
Dalbeth *et al*, 2014[27] (Dalbeth *et al* 2013[42])	Baseline (part 1)	139.8 (23.8)	–	410 (70)	2 (17)	2 (17) pts had ≥1 in 3 months
	6 months (part 1/2)	134.3 (24.3)	−5.5 (5.2)	440 (90)	2 (17)	0 (0) pts had ≥1 in 6 months
	6 months, 2 weeks (part 2)	NA	NA	510 (130)	NA	NA
	1.5 years (part 2)	100.3 (16.3)	−34 (11.0)	330 (90)	8 (67)	3 (25) pts had ≥1 in follow-up
Romero-Talamás *et al*, 2014[46]	I: Baseline	143‡	–	546 (120)	NA	20 (24) pts had ≥1 in 12 months
	I: 1 month	132‡	−11‡	NA	NA	NA (18) pts had ≥1 in 1 month
	I: 13 months	101‡	−31‡	336 (150)	NA	NA (8) pts had ≥1 in 12 months
	C: Baseline	106‡	–	462 (120)	NA	10 (18) pts had ≥1 in 12 months
	C: 1 month	105‡	−1‡	NA	NA	NA (2) pts had ≥1 in 1 month
	C: 13 months	104‡	−2‡	420 (96)	NA	NA (11) pts had ≥1 in 12 month
Zeng *et al*, 2012[49] *(In Chinese)*	I: Baseline	74.5 (3.50)	–	486 (41)	NA	All (100) had ≥1 in 6 months; 33 episodes
	I: 6 months	65.8 (4.44)	−8.7	420 (37)	NA	17 episodes; 48% fewer gout attacks
	C: Baseline	72.7 (3.26)	–	486 (41)	NA	All (100) had ≥1 in 6 months; 36 episodes
	C: 6 months	69.3 (7.78)	−3.4	467 (42)	NA	28 episodes; 22% fewer gout attacks
Perez-Ruiz *et al*, 2011[45] (*unpublished data*)	Lost weight: baseline	81.0 (11.0)	–	263 (59)§	23 (92)§	0 (0)§ pts with gout attacks at withdrawal
	Lost weight: mean 34 (26) months	77.8 (11.0)	−3.2 (5.6)	298 (59)§	27 (93)§	5 (17)§ pts with gout attacks during follow-up
	No weight loss: baseline	81.9 (8.6)	–	491 (95)§	150 (90)§	0 (0)§ pts with gout attacks at withdrawal
	No weight loss: mean 32 (28) months	84.3 (9.9)	2.4 (4.3)	509 (90)§	146 (90)§	80 (49)§ pts with gout attacks during follow-up
Zhu *et al*, 2013[29] (Sherwin *et al* 1981[47])	≤−10 kg: 7 years	NA	(range, ≤−10)	Change, −37 (−40 to −35)‡‡	OR 3.19 (1.99 to 5.09)**	NA
	−5 to −9.9 kg: 7 years	NA	(range, −9.9 to −5)	Change, −19 (−20 to −17)‡‡	OR 2.33 (1.75 to 3.11)**	NA
	−1 to −4.9 kg: 7 years	NA	(range, −4.9 to −1)	Change, −7 (−9 to −6)‡‡	OR 1.53 (1.24 to 1.89)**	NA
	No change: 7 years	NA	(range, −0.9 to 0.9)	Change, 0 (reference)‡‡	OR 1.0 (reference)**	NA
	+1 to +4.9 kg: 7 years	NA	(range, 1 to 4.9)	Change, 5 (4 to 7)‡‡	OR 1.01 (0.80 to 1.27)**	NA
	+5 to +9.9 kg: 7 years	NA	(range, 5 to 9.9)	Change, 17 (16 to 19)‡‡	OR 0.65 (0.45 to 0.95)**	NA
	≥+10 kg: 7 years	NA	(range, ≥10)	Change, 26 (23 to 29)‡‡	OR 0.58 (0.31 to 1.08)**	NA
Barskova *et al*, 2009[40] *(In Russian)*	Baseline	94‡	–	570 (110)	0 (0)	3†† (range, 1 to 6) pr. patient in 12 months
	6 months	91‡	−3‡	435 (91)	NA	NA
	12 months	91‡	−3‡	443 (107)	11 (48)	1†† (range, 0–2) pr. patient in 12 months
Friedman *et al*, 2008[43]	Baseline	143‡	–	NA	NA	NA
	6 months	NA	NA	NA	NA	7 (33%) pts had an attack during 6 months
Dessein *et al*, 2000[28] (Terkeltaub *et al* 2001[48])	Baseline	91.1†† (23.5)	–	570†† (100)	1 (8)††¶	2.1†† (SD, 0.8) pr. month (in 4 months)
	16 weeks	83.4†† (22.0)	−7.7†† (range: 0 to −21)	470†† (90)	7 (54)††¶	0.6†† (SD, 0.7) pr. month

Continued

Table 2 Continued

Author, year, (multiple publication)	Group: time point	Body weight, kg	Body weight change from baseline, kg	sUA, µmol/L	Achieving sUA target*, n pts	Gout attacks
Brandstetter *et al*, 1986[41] *(In German)*	I: Baseline	89.9 (10.9)	–	316 (46)	NA	NA
	I: 4 weeks	89.1 (10.6)	−0.8	368 (55)	NA	NA
	I: 3 months	87.3 (9.6)	−2.6	364 (29)	NA	NA
	I: 6 months	84.2 (8.4)	−5.7	320 (61)	NA	NA
	C: Baseline	77.0 (14.0)	–	297 (54)	NA	NA
	C: 4 weeks	77.0 (13.8)	0	304 (98)	NA	NA

The results are reported as means (SD) or number (%), unless otherwise indicated.
*Achieving sUA target, that is, sUA <360 µmol/L (6 mg/dL).
†Multivariable OR of recurrent gout attacks according to BMI change. Based on a conditional logistic regression adjusted for BMI, age, education, alcohol and coffee intake, presence of hypertension and diuretic use measured during the 12 months before the incident gout attack. It should be noted that non-overweight gout patients were included as well in the analysis.
‡Estimated from the BMI assuming a height of 1.70 m.
§Due to loss of data, the number of patients in the groups were 25 and 167 at baseline, and 29 and 163 at last follow-up.
¶Note that Dessein *et al* uses a cut-off of ≤510 µmol/L.
**Multivariable OR of achieving sUA target of ≤360 µmol/L according to the weight loss. Based on a conditional logistic regression adjusted for time-varying covariates, that is, age, congestive heart failure, hypertension, diuretic use, serum creatinine level, alcohol intake and dietary variables (intake of fructose, caffeine, total protein, polyunsaturated fat, monounsaturated fat, saturated fat and fibre). Reported with corresponding 95% CIs. It should be noted that non-overweight gout patients were included as well in their analysis.
††Median.
‡‡Based on linear mixed model adjusted for baseline covariates (race, education level and weight categories), and time-varying covariates, that is, age, congestive heart failure, hypertension, diuretic use, serum creatinine level, alcohol intake and dietary variables (intake of fructose, caffeine, total protein, polyunsaturated fat, monounsaturated fat, saturated fat and fibre). Reported with corresponding 95% CIs. It should be noted that non-overweight and non-gout patients were included as well in their analysis.
AEs, adverse events; BMI, body mass index; I, intervention group; C, control group; n, number; pts, patients; sUA, serum uric acid.

Table 3 Summary of findings and GRADE evidence profile

Quality assessment								No. of patients	Weight loss information		Effect		Quality
Studies with data	Study limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Other	Study	Weight loss/ No weight loss	Weight loss, kg (% of body weight)	Weight loss pr. month (% of body weight)	Effect on outcome	Overall conclusion	GRADE rating
Serum uric acid (µmol/L)													
7	Serious (–1)	Not serious	Not serious	Not serious	Not serious	Dose response (+1)	Nguyen	NA/NA	–	–	NA	Weight loss results in a decrease of sUA after medium/long follow-up.	Low ⊕⊕◯◯
							Dalbeth (part 1)	12/0	5.5 kg (3.9)	0.9 kg (0.7)	30* (–48 to 108)		
							Dalbeth (part 2)	12/0	34 kg (24.3)	2.8 kg (2.0)	–110* (–188 to –32)		
							Romero-Talamás	99/56	29 kg† (23)	2.4 kg (1.8)	–168 (–213 to –123)		
							Zeng	30‡/31	3.6 kg (4.8)	0.6 kg (0.8)	–47 (–66 to –27)		
							Perez-Ruiz	25§/167§	5.6 kg (6.9)	0.2 kg (0.2)	17 (-40 to 73)		
							Zhu	NA/NA	–	–	Dose–response relationship between weight change and sUA change¶		
							Barskova	23‡/0	3 kg† (3.2)	0.3 kg (0.3)	–127* (–183 to –71)		
							Friedman	21/0	NA	NA	NA		
							Dessein	13/0	7.7 kg** (8.4)	2.1 kg (2.3)	–100*†† (–174 to –26)		
							Brandstetter	11‡/11	5.7 kg (6.9)	1.0 kg (1.1)	–2 (–116 to 113)		
Achieving serum uric acid target that is, sUA<360 µmol/L													
4	Serious (–1)	Not serious	Not serious	Not serious	Not serious	Large effect (+1) Dose response (+1)	Nguyen	NA/NA	–	–	NA	Weight loss results in a higher chance of achieving sUA target after medium/ long follow-up.	Moderate ⊕⊕⊕◯
							Dalbeth (part 1)	12/0	5.5 kg (3.9)	0.9 kg (0.7)	0% (0/10) reduction in pts with raised sUA*		
							Dalbeth (part 2)	12/0	34 kg (24.3)	2.8 kg (2.0)	60% (6/10) reduction in pts with raised sUA*		
							Romero-Talamás	99/56	29 kg† (23)	2.4 kg (1.8)	NA		
							Zeng	30‡/31	3.6 kg (4.8)	0.6 kg (0.8)	NA		
							Perez-Ruiz	25§/167§	5.6 kg (6.9)	0.2 kg (0.2)	1% reduction in pts with raised sUA		
							Zhu	NA/NA	–	–	Dose–response relationship between weight change achieving serum uric acid target**		
							Barskova	23‡/0	3 kg† (3.2)	0.3 kg (0.3)	48% (11/23) reduction in pts with raised sUA*		
							Friedman	21/0	NA	NA	NA		
							Dessein	13/0	7.7 kg **(8.4)	2.1 kg (2.3)	50% (6/12) reduction in pts with raised sUA*‡‡		
							Brandstetter	11‡/11	5.7 kg (6.9)	1.0 kg (1.1)	NA		

Continued

Table 3 Continued

Quality assessment								No. of patients	Weight loss information		Effect		Quality
Studies with data	Study limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Other	Study	Weight loss/ No weight loss	Weight loss, kg (% of body weight)	Weight loss pr. month (% of body weight)	Effect on outcome	Overall conclusion	GRADE rating
Gout attacks													
8	Serious (–1)	Not serious	Not serious	Not serious	Not serious	Dose response (+1)	Nguyen	NA/NA	–	–	Dose–response relationship between BMI change and recurrent gout attacks**	Weight loss results in fewer gout attacks after medium/long follow-up.	Low ⊕⊕◯◯
							Dalbeth (part 1)	12/0	5.5 kg (3.9)	0.9 kg (0.7)	0 pts had ≥1 attack 6 months (At baseline two pts had ≥1 attack in 3 months)*		
							Dalbeth (part 2)	12/0	34 kg (24.3)	2.8 kg (2.0)	Three pts had ≥1 attack in 12 months (At baseline for part 2, 0 pts had ≥1 attack in 6 months)*		
							Romero-Talamás	99/56	29 kg† (23)	2.4 kg (1.8)	RR of 0.72 for ≥1 attack at follow-up		
							Zeng	30‡/31	3.6 kg (4.8)	0.6 kg (0.8)	39% fewer attacks at follow-up		
							Perez-Ruiz	25§/167§	5.6 kg (6.9)	0.2 kg (0.2)	RR of 0.35 for ≥1 attack at follow-up		
							Zhu	NA/NA	–	–	NA		
							Barskova	23‡/0	3 kg† (3.2)	0.3 kg (0.3)	33% fewer attacks*††		
							Friedman	21/0	NA	NA	33% (7/21) pts had attacks*§§		
							Dessein	13/0	7.7 kg** (8.4)	2.1 kg (2.3)	71% fewer attacks*††		
							Brandstetter	11‡/11	5.7 kg (6.9)	1.0 kg (1.1)	NA		

Modified from table made with GRADEpro (computer program on www.gradepro.org), McMaster University, 2014. Effect on outcome was calculated as change from baseline to latest follow-up for NRS with one group, difference in changes for NRS with two groups and difference between groups at follow-up for RCTs, unless otherwise indicated. Effects on outcomes are presented as means (95% CI), numbers or RRs.
*Studies with only one group, hence the effect is not a contrast between groups.
† Weight loss estimated from BMI.
‡ On average a weight loss was seen for these pts as a group. Hence, some individuals may not have lost weight.
§Due to loss of data in the study, the number of pts in the groups were 25 and 167 at baseline, and 29 and 163 at last follow-up, respectively. For calculation of 95% CI, a mean, that is, 26 and 165, was used.
¶It should be noted that non-overweight and non-gout patients were included as well.
**It should be noted that non-overweight gout patients were included as well.
††Based on medians.
‡‡Dessein[28] used a cut-off of ≤510 µmol/L.
§§For this study, no comparison (such as a comparison group or a measurement before the intervention) was provided. Hence, the number is the absolute number.
BMI, body mass index; GRADE, The Grading of Recommendations Assessment, Development and Evaluation; NRS, non-randomised study; pts, patients; RCT, randomised controlled trial; RR, risk ratio; sUA, serum uric acid.

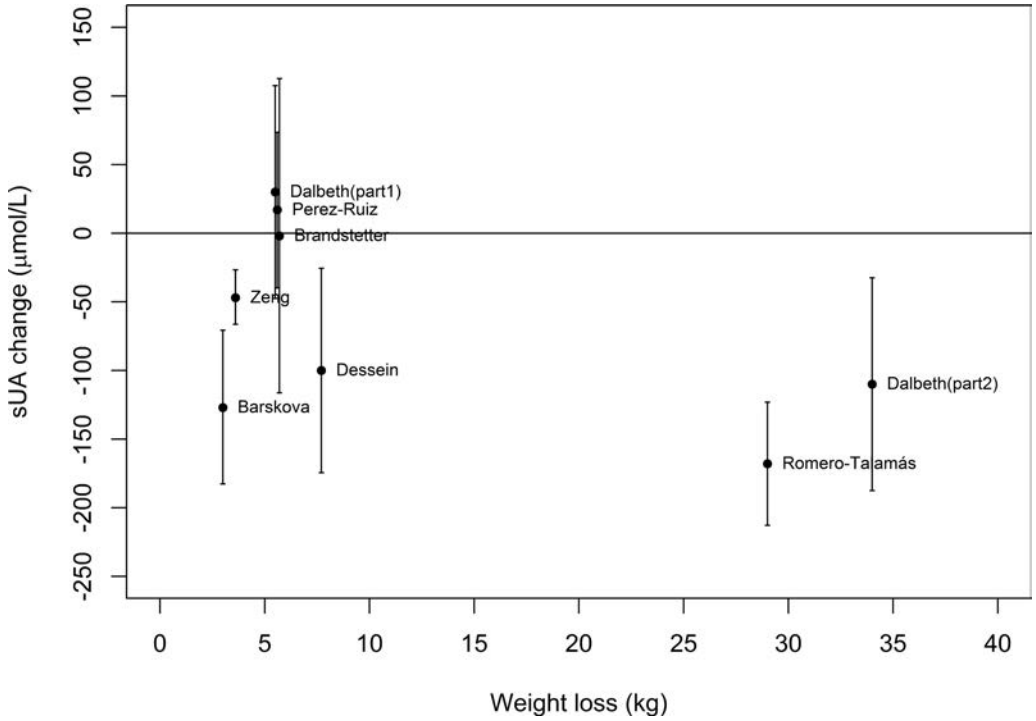


Figure 2 Relationship between weight loss and serum uric acid at latest follow-up. Estimates are shown with 95% confidence intervals. sUA, serum uric acid.

1% reduction in patients with raised sUA reported by Dalbeth *et al* (part 1)[27] and Perez-Ruiz *et al*,[45] respectively, are consistent with no change in sUA. Furthermore, achieving sUA target reported by Dessein *et al*[28] may be overestimated due to using a higher sUA cut-off of 510 µmol/L.

All studies, except two,[27 43] with data on gout attacks, showed a beneficial effect, and Dessein *et al*[28] reported 71% fewer attacks. Furthermore, a dose–response relationship was shown by Nguyen *et al*.[44] It should be noted that non-overweight patients were included in their analysis as well. Dalbeth *et al* (part 2)[27] reported an increase from zero patients experiencing ≥1 attack during 6 months to three patients during 12 months. This is possibly due to including the immediate postoperative phase where attacks could be a consequence of the increased sUA. Likewise, Romero-Talamás *et al*[46] report a possible increase from 24% experiencing ≥1 attack during 1 year at baseline to 18% during 1 month at 1-month follow-up, and a subsequently decrease to 8% during 1 year at last follow-up.

Only one study[40] included patients with tophi at baseline, therefore the impact on tophi could not be assessed.

Risk of bias assessment

The most frequent risk of bias was 'Bias due to confounding', with four studies rated critical due to studying one group without adjustment for confounders.[28 40 42 43] Five studies were rated serious (figure 3), and only the RCT[49] was rated low risk of bias. All studies were rated serious risk for 'Bias due to departures from intended interventions' (see possible confounders and cointerventions in online supplementary table S3). Reporting bias was suspected for two studies reporting change in BMI instead of change in weight.[40 46] Protocols were only found for two substudies[29 44] of one main study[39 47] and three studies reported no published protocol.[27 44 45]

Quality of the evidence using GRADE

For a beneficial effect of weight loss at medium-term/long-term follow-up, we evaluated the overall quality of evidence to be low for sUA, moderate for achieving sUA target and low for gout attacks.

DISCUSSION

Overall, we found low to moderate quality of evidence for beneficial effects of weight loss for overweight gout patients in terms of sUA, achieving sUA target and gout attacks. No or few data were available on our remaining prespecified outcomes. We did not find evidence for the optimal magnitude and intensity of weight loss. However, our data suggest that a weight loss of >7 kg and/or >2 kg per week from either surgery or diet results in a beneficial effect on sUA at medium-term/long-term follow-up based on three studies[27 28 46] and that weight loss of >3.5 kg showed beneficial effects on gout attacks at medium-term/long-term follow-up based on six studies.[27 28 40 45 46 49] However, with the present quality of evidence, further research may change these findings. WDdtAEs and SAEs were poorly reported. At short term, weight loss from bariatric surgery showed temporarily increased sUA levels and gout attacks, that is, a harmful effect, in the immediate postoperative period based on two studies.[27 46]

It is well known that there is a higher risk of gout attacks during the first months of urate-lowering therapy, postsurgery and starvation.[59] One hypothesis is that dramatic changes in sUA, rather than absolute level, triggers gout attacks.[60] In line with this, a study[61] comparing gout patients experiencing postoperative gout attacks with those who did not, find that the first group had higher presurgical sUA and a more rapid and larger decrease in sUA 3 days after surgery. Dalbeth *et al*[27] reported a drastic increase 2 weeks after surgery, which they suggested was due to renal dysfunction associated with major surgery, or metabolic effects from fasting or rapid weight loss (catabolic state),

Author, year and outcome	1. Bias due to confounding	2. Bias in selection of participants into the study	3. Bias in classification of interventions	4. Bias due to departures from intended interventions	5. Bias due to missing data	6. Bias in measurement of outcomes	7. Bias in selection of the reported result
Nguyen, 2016*							
Gout attacks	Serious	Moderate	Moderate	Serious	Moderate	Serious	Low
Dalbeth, 2014* (part 1)							
Physical function, SF36 PF	Critical	Serious	Moderate	Serious	Moderate	Serious	Low†
sUA and achieving sUA target	Critical	Serious	Moderate	Serious	Moderate	Low	Low
Body weight	Critical	Critical	Moderate	Serious	Moderate	Low	Low
Gout attacks	Critical	Serious	Moderate	Serious	Moderate	Serious	Low
SAEs	Critical	Serious	Moderate	Serious	Moderate	Serious	Low
Dalbeth, 2014* (part 2)							
Physical function, SF36 PF	Critical	Low	Low	Serious	Moderate	Serious	Low†
sUA and achieving sUA target	Critical	Low	Low	Serious	Moderate	Low	Low
Body weight	Critical	Low	Low	Serious	Moderate	Low	Low
Gout attacks	Critical	Low	Low	Serious	Moderate	Serious	Low
Romero-Talamás, 2014							
sUA	Serious	Low	Low	Serious	NI	Moderate	Moderate
Body weight (i.e. change in BMI)	Serious	Low	Low	Serious	NI	Moderate	Serious
Gout attacks	Serious	Low	Low	Serious	NI	Serious	Moderate
Zeng, 2012							
sUA	Low	Low	Low	Serious	Moderate	Low	Moderate
Body weight	Low	Low	Low	Serious	Moderate	Low	Moderate
Gout attacks	Low	Low	Low	Serious	Moderate	Serious	Moderate
Perez-Ruiz, 2011							
sUA and achieving sUA target	Serious	Low	Moderate	Serious	Low	Moderate	Low
Presence of tophi	Serious	Low	Moderate	Serious	Low	Low	Low
Body weight	Serious	Low	Moderate	Serious	Low	Low	Low
Gout attacks	Serious	Low	Moderate	Serious	Low	Serious	Low
Zhu, 2010*							
sUA and achieving sUA target	Serious	Moderate	Moderate	Serious	Moderate	Low	Moderate
Barskova, 2009							
sUA and achieving sUA target	Critical	Low	Low	Serious	Moderate	Low	Moderate
Body weight (i.e. change in BMI)	Critical	Low	Low	Serious	Moderate	Low	Serious
Gout attacks	Critical	Low	Low	Serious	Moderate	Serious	Moderate
WDdtAEs	Critical	Low	Low	Serious	Moderate	Serious	Moderate
Friedman, 2008							
Gout attacks	Critical	Low	Low	Serious	Low	Serious	Moderate
Dessein, 2000*							
sUA and achieving sUA target	Critical	Low	Low	Serious	Low	Low	Moderate
Body weight	Critical	Low	Low	Serious	Low	Serious	Moderate
Gout attacks	Critical	Low	Low	Serious	Low	Serious	Moderate
Brandstetter, 1986							
sUA	Serious	Low	Low	Serious	Low	Low	Moderate
Body weight	Serious	Low	Low	Serious	Low	Low	Moderate

Critical: Critical risk of bias is present. Serious: Serious risk of bias is present. Moderate: Moderate risk of bias is present. Low: Low risk of bias is present. NI: No information available.

Figure 3 Risk of bias summary figure. Similar outcomes has been put together in the figure but has been assessed separately. *Multiple publications existed. A primary publication was chosen. †Potentially serious risk of bias, since physical function was not reported in the article, but assessed low since data were provided from the author through email contact. BMI, body mass index; sUA, serum uric acid.

and they report one case of postoperative gout attack together with severe hyperuricaemia. Other factors increasing sUA levels are fasting,[62–64] dehydration[65] and tissue hypoxia.[66] Fasting-associated increase in sUA is likely due to tissue breakdown.[67 68] In line with this, daytime fasting during Ramadan, without weight loss, compared with non-fasting did not increase sUA or gout attacks in gout patients.[69]

Increased sUA seems to be related to decreased estimated glomerular filtration rate (eGFR).[70 71] This is probably related to sUA affecting blood pressure,[72] which may be caused by

increased vascular stiffness.[73 74] Reducing sUA may therefore have beneficial effect on susceptibility towards cardiovascular disease and diminished renal function.

In our study, we lacked evidence for many prespecified outcomes important to patients. Serum uric acid was among the most frequently reported outcomes and is recommended as a treatment target,[15–19 21–23] since elevated sUA is considered to cause the disease. Gout attacks in this study is not well defined and was reported in various ways and over various follow-up times. Therefore, stating fewer gout attacks following weight loss is not very specific and not necessarily assessable in smaller study sizes, or when attacks were not systematically assessed. At least three studies[27 43 45] did not point at reduced frequency of attacks, of which Friedman *et al*[43] did not report any baseline and Perez-Ruiz *et al*[45] did show less increase compared with control. Other studies can mask increasing number of attacks by reporting number of patients experiencing ≥1 attack over various follow-ups. Therefore, one could consider rating the evidence for gout attacks further down for indirectness.

Limitations of our methods include no independent double study selection, data extraction or risk of bias assessment. A limitation of investigating weight loss per se is that weight loss can be a consequence of many different interventions, that is, cointerventions, or conditions. Hence, it was impossible to ensure the weight reduction to be the only difference in terms of intervention from the comparison group, resulting in the inclusion of a wide variety of study settings. This is also observed as for which variables have been measured longitudinally. The included cohort studies stratifying according to weight loss may include unintentional weight loss for example, from illness, which is not relevant as intervention. Adding this to the fact that the majority of our included studies did not have a comparison group introducing non-controllable confounding, we cannot be sure that weight loss is accountable for all the effects observed. As a result, the implementation of weight loss intervention in clinical practice cannot be specified from the included studies. Taking the limitations of the available evidence into account, one may suggest, in order to address the effect of weight loss on sUA, that there currently is a need to perform a systematic review and meta-analysis of data not only for gout patients.

In conclusion, the available evidence is in favour of weight loss for overweight gout patients at medium-term/long-term follow-up on sUA, achieving sUA target and gout attacks. However, the evidence is of low, moderate and low quality, respectively. Harms were poorly reported. However, gout attacks might occur at short term when initiating treatment. We believe that there is an urgent need to initiate rigorous prospective studies (preferably RCTs) to provide more trustworthy estimates of gout-related benefits and harms including the effect on joint pain, tophi, physical function, HRQoL, adverse events and patient global assessment. Future research should aim at identifying the optimal magnitude and intensity of weight loss, the preferred method of weight loss, including prevention of flare, which cointerventions result in a better effect, and which gout patients will benefit the most, for example, grouped according to type (and possibly severity) of overweight and comorbidities.

Author affiliations
[1]The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark
[2]Department of Physical and Occupational Therapy, Bispebjerg and Frederiksberg, Copenhagen, Denmark
[3]The Research Initiative for Activity Studies and Occupational Therapy, General Practice, Department of Public Health, University of Southern Denmark, Odense, Denmark
[4]Department of Nutrition, Exercise and Sports, Faculty of Science, University of Copenhagen, Copenhagen, Denmark
[5]Center for Diabetes Research, Gentofte Hospital, University of Copenhagen, Hellerup, Denmark
[6]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[7]NNF Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[8]Institutode Salud Musculoesquelética, Madrid, Spain
[9]Department of Medicine, University of Otago, Wellington, New Zealand
[10]Department of Medicine, University of Alabama at Birmingham, & Birmingham Veterans Affairs Medical Center, Birmingham, Alabama, USA
[11]Rheumatology Division, Hospital de Cruces, Baracaldo, Spain

Acknowledgements The authors would like to thank the Copenhagen University Library, Frederiksberg for their hard work in retrieving full texts from all over the world. In particular, we want to thank the librarian, Karen Bendix Larsen. We would also like to thank for assistance with translation of articles in Russian, Chinese and Bulgarian from Natalia Manilo, MD, Department of Rheumatology, Rigshospitalet Glostrup and Frederiksberg, Copenhagen, Denmark; Tao Ma, MD, PhD, Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of Copenhagen, Copenhagen, Denmark; and Nora Vladimirova, MD, Department of Rheumatology, Rigshospitalet Glostrup and Frederiksberg, Copenhagen, Denmark, respectively. Furthermore, we would like to thank professor Nicola Dalbeth, Department of Rheumatology, Counties Manukau District Health Board, Auckland, New Zealand, who responded to our data and information requests.

Contributors Study concept and design: SMN, EMB, LEK and RC. Drafting of the manuscript: SMN, EMB and RC. Search strategy: EMB and SMN. Study selection, data extraction, bias assessment and synthesis: SMN, EMB and RC. Critical revision of the manuscript for important intellectual content and final approval before submission: All authors. Obtained funding: HB, LEK and RC.

Funding The Parker Institute, Bispebjerg and Frederiksberg Hospital is supported by a core grant from the Oak Foundation (OCAY-13-309). This research received a specific grant from the will of Mrs Elise Fredriksen; the Oak Foundation had no role in study design or writing of this manuscript.

Competing interests This study had no financial competing interests. The Parker Institute is grateful for the financial support received from public and private foundations, companies and private individuals over the years. The Oak Foundation is a group of philanthropic organisations that, since its establishment in 1983, has given grants to not-for-profit organisations around the world.

Patient consent No patients were directly included in the study (only in the primary studies of this review).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Smith EU, Díaz-Torné C, Perez-Ruiz F, *et al*. Epidemiology of gout: an update. *Best Pract Res Clin Rheumatol* 2010;24:811–27.
2 Choi HK, Mount DB, Reginato AM. Pathogenesis of gout. *Ann Intern Med* 2005;143:499–516.
3 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.
4 Cassetta M, Gorevic PD. Crystal arthritis. Gout and pseudogout in the geriatric patient. *Geriatrics* 2004;59:25–30.
5 Masseoud D, Rott K, Liu-Bryan R, *et al*. Overview of hyperuricaemia and gout. *Curr Pharm Des* 2005;11:4117–24.
6 Dalbeth N, Fransen J, Jansen TL, *et al*. New classification criteria for gout: a framework for progress. *Rheumatology* 2013;52:1748–53.
7 Neogi T, Jansen TL, Dalbeth N, *et al*. 2015 Gout classification criteria: an American College of Rheumatology/European League against Rheumatism collaborative initiative. *Ann Rheum Dis* 2015;74:1789–98.
8 Richette P, Bardin T. Gout. *Lancet* 2010;375:318–28.

9 Bernal JA, Quilis N, Andrés M, *et al*. Gout: optimizing treatment to achieve a disease cure. *Ther Adv Chronic Dis* 2016;7:135–44.
10 Grassi D, Pontremoli R, Bocale R, *et al*. Therapeutic approaches to chronic hyperuricemia and gout. *High Blood Press Cardiovasc Prev* 2014;21:243–50.
11 Kiltz U, Smolen J, Bardin T, *et al*. Treat-to-target (T2T) recommendations for gout. *Ann Rheum Dis* 2016;76:1–7.
12 Chu NF, Wang DJ, Liou SH, *et al*. Relationship between hyperuricemia and other cardiovascular disease risk factors among adult males in Taiwan. *Eur J Epidemiol* 2000;16:13–17.
13 Wang H, Wang L, Xie R, *et al*. Association of serum uric acid with body Mass Index: a Cross-Sectional Study from Jiangsu Province, China. *Iran J Public Health* 2014;43:1503–9.
14 Qaseem A, Harris RP, Forciea MA, *et al*. And recurrent gout: a clinical Practice Guideline from the American College of Physicians. *Ann Intern Med* 2016:1–11.
15 Richette P, Doherty M, Pascual E, *et al*. 2016 updated EULAR evidence-based recommendations for the management of gout. *Ann Rheum Dis* 2017;76:29–42.
16 Hamburger M, Baraf HSB, Adamson TC, *et al*. 2011 recommendations for the diagnosis and management of gout and hyperuricemia. *Phys Sportsmed* 2011;39:98–123.
17 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* 2007;46:1372–4.
18 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* 2012;64:1431–46.
19 Manara M, Bortoluzzi A, Favero M, *et al*. Italian society of rheumatology recommendations for the management of gout. *Reumatismo* 2013;65:4–21.
20 Romeijnders AC, Gorter KJ. [Summary of the Dutch College of General Practitioners' "Gout" Standard]. *Ned Tijdschr Geneeskd* 2002;146:309–13.
21 Sivera F, Andrés M, Carmona L, *et al*. Multinational evidence-based recommendations for the diagnosis and management of gout: integrating systematic literature review and expert opinion of a broad panel of rheumatologists in the 3e initiative. *Ann Rheum Dis* 2014;73:328–35.
22 Yamanaka H. Japanese guideline for the management of hyperuricemia and gout: second edition. *Nucleosides Nucleotides Nucleic Acids* 2011;30:1018–29.
23 Zhang W, Doherty M, Bardin T, *et al*. EULAR evidence based recommendations for gout. Part II: Management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2006;65:1312–24.
24 Maglio C, Peltonen M, Neovius M, *et al*. Effects of bariatric surgery on gout incidence in the Swedish Obese Subjects study: a non-randomised, prospective, controlled intervention trial. *Ann Rheum Dis* 2016;76:1.
25 Tsunoda S, Kamide K, Minami J, *et al*. Decreases in serum uric acid by amelioration of insulin resistance in overweight hypertensive patients: effect of a low-energy diet and an insulin-sensitizing agent. *Am J Hypertens* 2002;15:697–701.
26 Richette P, Poitou C, Manivet P, *et al*. Weight loss, Xanthine Oxidase, and serum urate levels: a Prospective Longitudinal Study of Obese Patients. *Arthritis Care Res* 2016;68:1036–42.
27 Dalbeth N, Chen P, White M, *et al*. Impact of bariatric surgery on serum urate targets in people with morbid obesity and diabetes: a prospective longitudinal study. *Ann Rheum Dis* 2014;73:797–802.
28 Dessein PH, Shipton EA, Stanwix AE, *et al*. Beneficial effects of weight loss associated with moderate calorie/carbohydrate restriction, and increased proportional intake of protein and unsaturated fat on serum urate and lipoprotein levels in gout: a pilot study. *Ann Rheum Dis* 2000;59:539–43.
29 Zhu Y, Zhang Y, Choi HK. The serum urate-lowering impact of weight loss among men with a high cardiovascular risk profile: the multiple risk factor intervention trial. *Rheumatology* 2010;49:2391–9.
30 Shamseer L, Moher D, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
31 WHO. BMI classification. WHO Globalglobal Databasedatabase on Body Mass Index. http://apps.who.int/bmi/index.jsp?introPage=intro_3.html (accessed 22 Mar 2016).
32 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.
33 Schumacher HR, Taylor W, Edwards L, *et al*. Outcome domains for studies of acute and chronic gout. *J Rheumatol* 2009;36:2342–5.
34 Singh JA, Taylor WJ, Simon LS, *et al*. Patient-reported outcomes in chronic gout: a report from OMERACT 10. *J Rheumatol* 2011;38:1452–7.
35 Sterne JA, Hernán MA, Reeves BC, *et al*. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.
36 Sterne JAC, Higgins JPT, Reeves BC. On behalf of the development group for ACROBAT- NRSI. A cochrane risk Of bias assessment tool: for non-randomized studies of interventions. 2014 http://www.riskofbias.info
37 Dwan K, Gamble C, Kolamunnage-Dona R, *et al*. Assessing the potential for outcome reporting bias in a review: a tutorial. *Trials* 2010;11:52.
38 Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
39 Multiple risk factor intervention Trial Research Group. Multiple risk factor intervention trial. risk factor changes and mortality results. *JAMA* 1982;248:1465–77.
40 Barskova VG, Eliseev MS, Kudaeva FM, *et al*. Effect of metformin on the clinical course of gout and insulin resistance [title translated from russian]. *Klin Med* 2009;87:41–6.
41 Brandstetter G, Hoffmann H, Maderbacher H, *et al*. Urikosurische Wirkung eines Neuen Betarezeptorenblocker-Diuretikum-Kombinationspräparates. *Acta Med Austriaca* 1986;13:29–37.
42 Dalbeth N, Chen P, White M, *et al*. Impact of bariatric surgery on serum urate targets in people with morbid obesity and diabetes: a prospective longitudinal study. *Ann Rheum Dis* 2014;73:797–802.
43 Friedman JE, Dallal RM, Lord JL. Gouty attacks occur frequently in postoperative gastric bypass patients. *Surg Obes Relat Dis* 2008;4:11–13.
44 Nguyen UD, Zhang Y, Louie-Gao Q, *et al*. Obesity obesity paradox in recurrent attacks of gout in observational studies: clarification and remedy. *Arthritis Care Res* 2017;69:561-566.
45 Perez-Ruiz F, Herrero-Beites AM, Carmona L. A two-stage approach to the treatment of hyperuricemia in gout: the "dirty dish" hypothesis. *Arthritis Rheum* 2011;63:4002–6.
46 Romero-Talamás H, Daigle CR, Aminian A, *et al*. The effect of bariatric surgery on gout: a comparative study. *Surg Obes Relat Dis* 2014;10:1161–5.
47 Sherwin R, Kaelber CT, Kezdi P, *et al*. The multiple risk factor intervention trial (MRFIT) II. The development of the protocol. *Prev Med* 1981;10:402–25.
48 Terkeltaub R. Syndrome X and gout: benefits of altered diet. *Curr Rheumatol Rep* 2001;3:9–10.
49 Zeng YC, Huang SF, GP M, *et al*. Effects of adjusted proportional macronutrient intake on serumuric acid, blood lipids, renal function, and outcome of patients with gout and overweight[translated from Chinese]. *Chinese J Clin Nut* 2012;20:210–4.
50 Eliseev MS, Barskova VG, Denisov IS. [Time course of changes in the clinical manifestations of gout in men: data of a 7-year retrospective follow-up]. *Ter Arkh* 2015;87:10–15.
51 Se En AO. Efficacy and safety of restrictive bariatric procedure in class I obesity population. *Obes Surg* 2013;23:843.
52 Kreider R, Oliver JM, Kresta JY, *et al*. Effects of diet type during an exercise and weight loss program on markers of metabolic syndrome in women with elevated uric acid levels. *Faseb J* 2011:25.
53 Lu N, Shai I, Zhang Y, *et al*. High-protein diet (Atkins Diet) and uric acid response. *Arthritis and rheumatism* 2014;66:S71–2.
54 Masuo K, Kawaguchi H, Mikami H, *et al*. Changes in serum uric acid, sympathetic activity, plasma insulin, and blood pressure levels during weight loss. *J Hypertens* 2003;21:S328–9.
55 Masuo K, Lambert GW. Effects of weight loss on serum uric acid concentrations. *Circulation* 2013;127:1.
56 Tinahones FJ, Soriguer FJ, Collantes E, *et al*. Decreased triglyceride levels with low calorie diet and increased renal excretion of uric acid in hyperuricaemic-hyperlipidaemic patients. *Ann Rheum Dis* 1995;54:609–10.
57 Department of Molecular, Endocrinology Metabolism, Graduate School of Medical Dental Sciences Tokyo, Medical Dental, University. Effect of febuxostat on vascular endothelial function in patients with hyperuricemia. in: umin clinical trials registry (UMIN-CTR) [Internet]. 1989 https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr.cgi?function=brows&action=brows&type=summary&language=E&recptno=R000010441 (accessed 26 apr 2016).
58 Iwatani M. Diet therapy for management of hyperuricemia and gout [title translated from Japanese]. *Nippon Rinsho* 2003;61(Suppl 1):184–92.
59 Richette P, Bardin T. Purine-rich foods: an innocent bystander of gout attacks? *Ann Rheum Dis* 2012;71:1435–6.
60 Terkeltaub R. Pathogenesis of Monosodium Urate Crystal-Induced Inflammation. In: Gresser U, Zöllner N, eds. *Urate deposition in man and its clinical consequences. Berlin, Heidelberg: springer Berlin Heidelberg*, 1991:97–106.
61 Kang EH, Lee EY, Lee YJ, *et al*. Clinical features and risk factors of postsurgical gout. *Ann Rheum Dis* 2008;67:1271–5.
62 Ogryzlo MA. Hyperuricemia induced by high fat diets and starvation. *Arthritis Rheum* 1965;8:799–822.
63 Maclachlan MJ, Rodnan GP. Effect of food, fast and alcohol on serum uric acid and acute attacks of gout. *Am J Med* 1967;42:38–57.
64 Drenick EJ, Hyperuricemia DEJ. Hyperuricemia, acute gout, renal insufficiency and urate nephrolithiasis due to starvation. *Arthritis Rheum* 1965;8:988–97.
65 Feinstein EI, Quion-Verde H, Kaptein EM, *et al*. Severe hyperuricemia in patients with volume depletion. *Am J Nephrol* 1984;4:77–80.
66 Woolliscroft JO, Colfer H, Fox IH. Hyperuricemia in acute illness: a poor prognostic sign. *Am J Med* 1982;72:58–62.
67 de Oliveira EP, Burini RC. High plasma uric acid concentration: causes and consequences. *Diabetol Metab Syndr* 2012;4:12.
68 Rock KL, Kataoka H, Lai JJ. Uric acid as a danger signal in gout and its comorbidities. *Nat Rev Rheumatol* 2013;9:13–23.
69 Habib G, Badarny S, Khreish M, *et al*. The impact of Ramadan fast on patients with gout. *J Clin Rheumatol* 2014;20:353–6.

70 Koratala A, Singhania G, Alquadan KF, *et al*. Serum uric acid exhibits inverse relationship with estimated glomerular Filtration Rate. *Nephron* 2016;134:231–7.
71 Kuwabara M, Bjornstad P, Hisatome I, *et al*. Elevated serum uric acid Level predicts rapid decline in kidney function. *Am J Nephrol* 2017;45:330–7.
72 Sidoti A, Nigrelli S, Rosati A, *et al*. Body mass index, fat free mass, uric acid, and renal function as blood pressure levels determinants in young adults. *Nephrology* 2017;22:279–85.
73 Kuwabara M, Niwa K, Hisatome I, *et al*. Asymptomatic hyperuricemia without comorbidities predicts cardiometabolic diseases: five-year japanese cohort study. *Hypertension* 2017;69.
74 Mehta T, Nuccio E, McFann K, *et al*. Association of Uric Acid with vascular stiffness in the Framingham Heart Study. *Am J Hypertens* 2015;28:877–83.
75 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264–9w64.

EXTENDED REPORT

Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome

Esther Mossel,[1] Konstantina Delli,[2] Jolien F van Nimwegen,[1] Alja J Stel,[1] Frans G M Kroese,[1] Fred K L Spijkervet,[2] Arjan Vissink,[2] Suzanne Arends,[1] Hendrika Bootsma,[1] on behalf of the EULAR US-pSS Study Group

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-2017-211250).

[1]Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[2]Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Correspondence to
Professor Hendrika Bootsma, Department of Rheumatology and Clinical Immunology, University Medical Center Groningen Hanzeplein 1, 9713 GZ, Groningen The Netherland; h.bootsma@umcg.nl

EM and KD contributed equally.

Received 2 February 2017
Revised 30 May 2017
Accepted 1 July 2017
Published Online First 28 July 2017

ABSTRACT

Objective To assess the validity of ultrasound of major salivary glands (sUS) compared with parotid and labial gland biopsies, sialometry, anti-SSA/Ro antibody status and classification criteria in patients clinically suspected with primary Sjögren's syndrome (pSS).

Methods 103 consecutive outpatients with clinically suspected pSS underwent sUS. Parenchymal echogenicity, homogeneity, hypoechogenic areas, hyperechogenic reflections and clearness of salivary gland border were scored according to the Hocevar scoring system. Total ultrasound score was calculated as the sum of these domains (range 0–48).

Results Absolute agreement between sUS and parotid (83%) and labial (79%) gland biopsy outcome was good. Negative sUS predicts negative parotid gland biopsy, and positive sUS predicts positive labial gland biopsy. Compared with the American European Consensus Group (AECG) classification, sUS showed an absolute agreement of 82%, sensitivity of 71% and specificity of 92%. Compared with the American College of Rheumatology (ACR) classification, absolute agreement was 86%, sensitivity was 77% and specificity was 92%. Compared with the ACR-European League Against Rheumatism (EULAR) classification, absolute agreement was 80%, sensitivity was 67% and specificity was 94%. Positive sUS predicts classification, but negative sUS does not exclude classification. The combination of positive sUS with presence of anti-SSA/Ro antibodies or negative sUS with absence of anti-SSA/Ro antibodies showed a high predictive value for classification as pSS or non-pSS.

Conclusion In our prospective inception cohort study derived from daily clinical practice, absolute agreement between sUS and salivary gland biopsies was slightly higher for parotid compared with labial gland biopsies. The combination of positive sUS and presence of anti-SSA/Ro antibodies highly predicts classification according to the AECG, ACR and ACR-EULAR classification criteria.

To cite: Mossel E, Delli K, van Nimwegen JF, *et al*. *Ann Rheum Dis* 2017;**76**:1883–1889.

INTRODUCTION

Primary Sjögren's syndrome (pSS) is a chronic, systemic autoimmune disease characterised by inflammation of the exocrine glands, with an estimated prevalence of 0.05% in the general population.[1] Most patients with pSS suffer from xerostomia, keratoconjunctivitis sicca and extreme fatigue.[2] In addition, different extraglandular manifestations, most frequently arthralgia, arthritis and myalgia, may be present.[2]

Currently, multiple criteria sets are available for the classification of pSS. In 2002, the American European Consensus Group (AECG) criteria were developed, and although not endorsed by the American College of Rheumatology (ACR) or European League Against Rheumatism (EULAR) these are yet most commonly used in daily clinical practice.[3 4] In 2012, the ACR criteria were developed and provisionally approved by the ACR,[5] but these criteria were not well received by many Sjögren's syndrome (SS) experts.[6] In order to develop international consensus on classification criteria, the ACR-EULAR criteria were recently introduced, endorsed by both EULAR and ACR.[4 7] In all three sets, salivary gland biopsies and presence of anti-SSA/Ro antibodies play a significant role in classifying patients as pSS.[3–5] In the AECG and ACR-EULAR criteria, salivary gland involvement is also assessed by unstimulated whole saliva flow (UWS).[3 4]

Ultrasound of major salivary glands (sUS) is an upcoming diagnostic method to assess involvement of major salivary glands in pSS.[8 9] sUS is well tolerated, non-invasive, inexpensive, non-irradiating and widely available in the rheumatological outpatient clinics, but its reliability depends greatly on its operator. A recent meta-analysis assessing the diagnostic properties of sUS in pSS reported a pooled sensitivity of 69% and specificity of 92%. This meta-analysis also revealed a large clinical and methodological heterogeneity between studies, which hampered interpretation of pooled outcomes and influenced the results reported in the various studies.[8] Thus, the possible role of sUS in the diagnosis of pSS remains unclear.[8–10]

This study assesses the validity of sUS compared with parotid and labial gland biopsies, sialometry, anti-SSA/Ro antibody status and classification criteria in patients clinically suspected with pSS.

MATERIALS AND METHODS

Patients

The present cross-sectional study is based on prospective data from the multidisciplinary Sjögren's expertise centre in the University Medical

Center Groningen, a tertiary referral centre. One hundred and ten consecutive patients clinically suspected with pSS, who underwent sUS as part of the diagnostic work-up and were over 18 years of age, were included. All patients underwent the diagnostic work-up. All domains of the AECG, ACR and ACR-EULAR criteria were assessed, including parotid and/or labial gland biopsy.

Primary assessment

Ultrasonography

All patients were examined with the same ultrasonographic scanner (Esaote MyLab Seven, Genova, Italy), equipped with a high-resolution linear scanner (4-13 MHz). Patients were examined in supine position with their neck slightly extended and turned away from the examined side.[11 12] The Hocevar *et al*[12] scoring system was used investigating (1) parenchymal echogenicity compared with the thyroid gland, graded 0–1; (2) homogeneity, graded 0–3; (3) presence of hypoechogenic areas, graded 0–3; (4) hyperechogenic reflections, graded 0–3 in parotid glands and 0–1 in submandibular glands; and (5) clearness of the salivary gland border, graded 0–3, in both parotid and submandibular salivary glands. Total ultrasound score was the sum of these five domains and can range from 0 to 48.[12]

Other assessments

Parotid and/or labial gland biopsies were considered positive if the focus score (defined as the number of mononuclear infiltrates containing ≥50 lymphocytes/4 mm^2 of glandular tissue) was ≥1.[13–15] UWS was evaluated by measuring the saliva production in 15 min.[3] UWS ≤1.5 mL/15 min was considered abnormal.[16] Serum levels of anti-SSA/Ro and anti-SSB/La antibodies were assessed with ELISA tests.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics V.23. Descriptive parameters were expressed as number of patients (%) or mean (SD).

Receiver operating characteristic analysis was performed to determine the accuracy of sUS to predict parotid or labial gland biopsies, UWS, anti-SSA/Ro antibody status and classification as pSS. Area under the curve (AUC) was interpreted as no discrimination (0–0.5), poor accuracy (0.5–0.7), fair (0.7–0.8), good (0.8–0.9) or excellent (0.9–1.0).[17] The optimal cut-off point for sUS positivity was determined according to the highest combination of sensitivity and specificity.

The percentage of absolute agreement between sUS outcome and parotid or labial gland biopsies, UWS, anti-SSA/Ro antibody status and classification according to the classification criteria was determined. The association between ultrasound and UWS was analysed using Spearman correlation coefficient (ρ), and interpreted as poor agreement (0.0–0.2), fair (0.2–0.4), moderate (0.4–0.6), good (0.6–0.8) or excellent (0.8–1.0).[18] Furthermore, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated.

Mann-Whitney U tests were used to evaluate differences in total ultrasound score between patients with (1) positive versus negative parotid or labial gland biopsies, (2) UWS ≤1.5 mL/15 min versus UWS >1.5 mL/15 min, (3) presence versus absence of anti-SSA/Ro and/or anti-SSB/La antibodies and (4) pSS versus non-pSS according to the classification criteria. p Values <0.05 were considered statistically significant.

RESULTS

A flow chart of inclusion and exclusion of patients and information about the number of patients included in the analyses on salivary gland biopsies, sialometry and anti-SSA/Ro antibody status is presented in figure 1. Of the 103 included patients, the mean age was 50 years (15), 90% were female and the mean total ultrasound score was 15 (10) (see online supplementary table S1). For research purposes, 43 patients underwent parotid and labial gland biopsies. Of the remaining patients, 35 had a parotid gland biopsy only and 13 had a labial gland biopsy only.

Ultrasound versus salivary gland biopsy

The accuracy of sUS to predict a parotid gland biopsy outcome was good, with an AUC of 0.849 (95% CI 0.746 to 0.952) and optimal cut-off point of 15 (see online supplementary table S2). The absolute agreement between sUS outcome and parotid gland biopsy was 83% (65/78), with a sensitivity of 75% (21/28), specificity of 88% (44/50), PPV of 78% (21/27) and NPV of 86% (44/51) (table 1).

The accuracy of sUS to predict a labial gland biopsy outcome was good, with an AUC of 0.824 (95% CI 0.714 to 0.934) and optimal cut-off point of 14 (see online supplementary table S2).

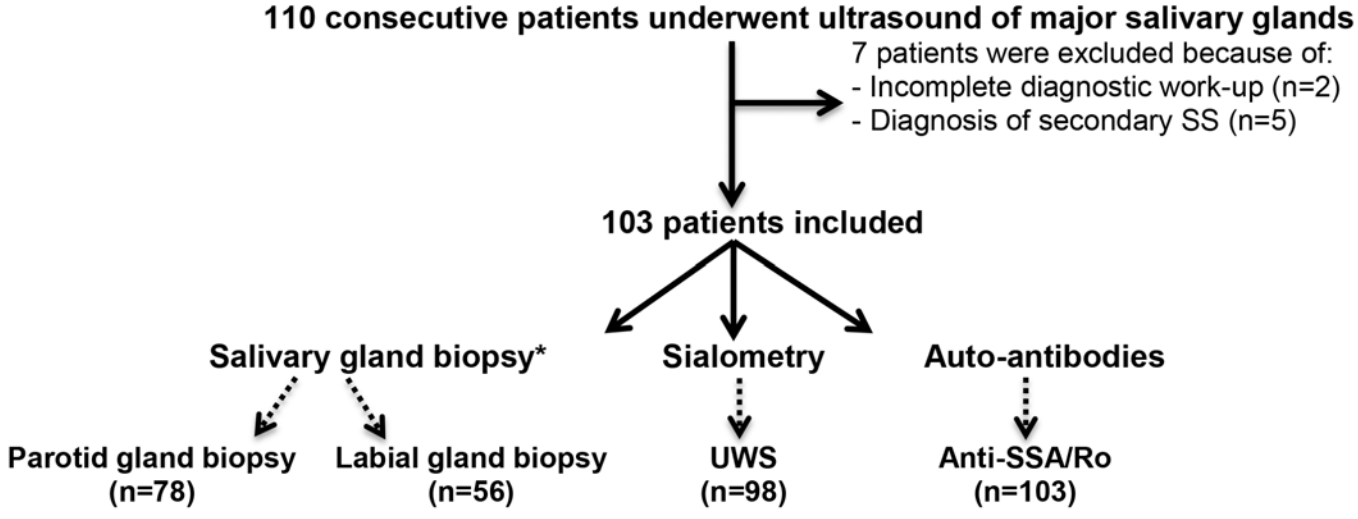


Figure 1 Flow chart of number of patients with available data per analysis. *43 patients underwent parotid gland as well as labial gland biopsy. SS, Sjögren's syndrome; UWS, unstimulated whole saliva.

Table 1 Ultrasound of major salivary glands versus salivary gland biopsy, sialometry and anti-SSA/Ro antibodies status

	Parotid gland biopsy (n=78)	Labial gland biopsy (n=56)	UWS (n=98)	Anti-SSA/Ro antibodies (n=103)
Optimal cut-off point	15	14	15	15
% Absolute agreement	83.3% (65/78)	78.6% (44/56)	66.3% (65/98)	81.6% (84/103)
Sensitivity	75.0% (21/28)	72.4% (21/29)	54.5% (30/55)	69.8% (37/53)
Specificity	88.0% (44/50)	85.2% (23/27)	81.4% (35/43)	94.0% (47/50)
PPV	77.8% (21/27)	84.0% (21/25)	78.9% (30/38)	92.5% (37/40)
NPV	86.3% (44/51)	74.2% (23/31)	58.3% (35/60)	74.6% (47/63)

NPV, negative predictive value; PPV, positive predictive value; UWS, unstimulated whole saliva.

The absolute agreement between sUS outcome and labial gland biopsy was 79% (44/56), with a sensitivity of 72% (21/29), specificity of 85% (23/27), PPV of 84% (21/25) and NPV of 74% (23/31) (table 1).

Total ultrasound score was significantly higher in patients with positive parotid or labial gland biopsies compared with patients with negative parotid or labial gland biopsies ($p<0.001$; figure 2).

Ultrasound versus sialometry

The accuracy of sUS to predict UWS outcome was poor, with AUC of 0.696 (95% CI 0.593 to 0.799). The absolute agreement between sUS outcome and UWS was 66% (65/98) (table 1).

Total ultrasound score was significantly higher in patients with UWS ≤1.5 mL/15 min compared with patients with UWS >1.5 mL/15 min ($p<0.001$; see online supplementary figure S1A). There was a fair reversed association between total ultrasound score and UWS total flow ($\rho=-0.366$) (see online supplementary figure S2B).

Ultrasound versus autoantibodies

The accuracy of sUS to predict anti-SSA/Ro antibody status was good, with an AUC of 0.803 (95% CI 0.711 to 0.894). The absolute agreement between sUS outcome and anti-SSA/Ro antibody status was 82% (84/103) (table 1).

Total ultrasound score was significantly higher in patients with anti-SSA antibodies compared with patients without anti-SSA/Ro antibodies ($p<0.001$; see online supplementary figure 2).

Ultrasound versus classification criteria

For the following analyses, sUS is compared with classification criteria with the outcome of parotid gland biopsy considered as an item of these criteria. For the comparison of sUS with classification criteria with the outcome of labial gland biopsy considered as an item, see table 2, see online supplementary table S3 and figure 3.

The accuracy of sUS to predict AECG classification was good, with an AUC of 0.826 (95% CI 0.735 to 0.918) and optimal cut-off point of 15 (see online supplementary table S3). The absolute agreement between sUS outcome and AECG classification was 82% (80/97), with a sensitivity of 71% (32/45), specificity of 92% (48/52), PPV of 89% (32/36) and NPV of 79% (48/61) (table 2).

The accuracy of sUS to predict ACR classification was good, with an AUC of 0.862 (95% CI 0.777 to 0.947) and optimal cut-off point of 15 (see online supplementary table S3). The absolute agreement between sUS outcome and ACR classification was 86% (83/97), with a sensitivity of 77% (34/44), specificity of 92% (49/53), PPV of 89% (34/38) and NPV of 83% (49/59) (table 2).

The accuracy of sUS to predict ACR-EULAR classification was good, with an AUC of 0.802 (95% CI 0.710 to 0.894) and optimal cut-off point of 15 (see online supplementary table S3). The absolute agreement between sUS outcome and ACR-EULAR classification was 80% (79/99), with a sensitivity of 67% (35/52), specificity of 94% (44/47), PPV of 92% (35/38) and NPV of 72% (44/61) (table 2).

Total ultrasound score was significantly higher in pSS versus non-pSS according to the classification criteria ($p<0.001$; figure 3).

Predictive value of the combination of sUS and antibody status

In patients with positive sUS combined with anti-SSA/Ro antibodies, 78% (14/18) had a positive parotid gland biopsy and 94% (17/18) had a positive labial gland biopsy.

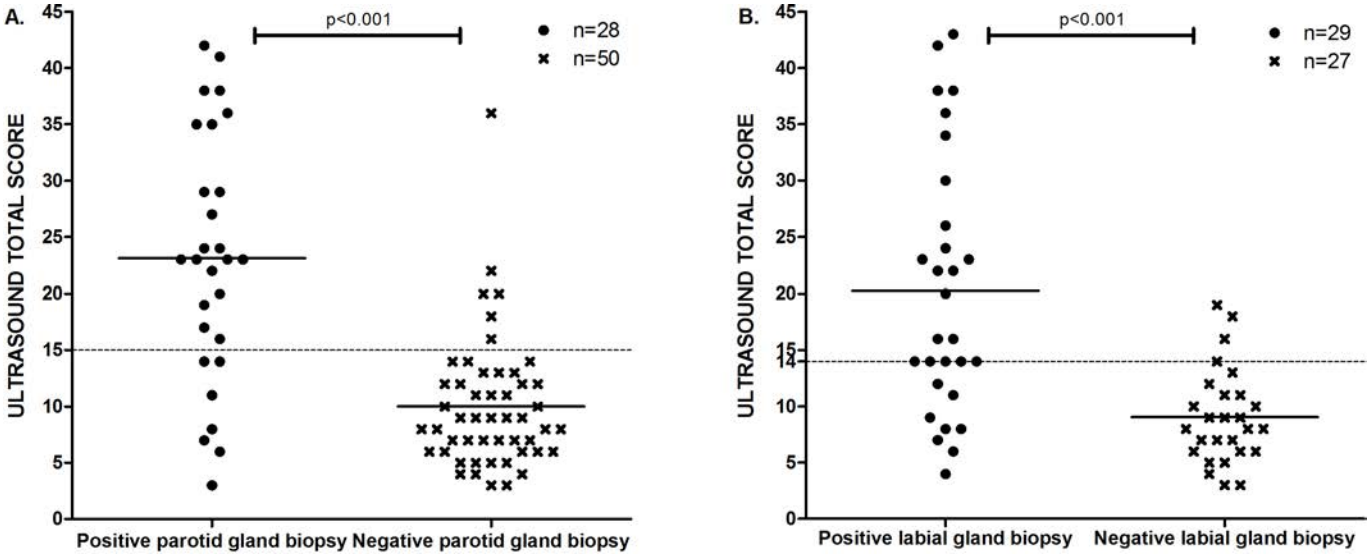


Figure 2 Ultrasound total score compared with salivary gland biopsy. (A) Positive versus negative parotid gland biopsy. (B) Positive versus negative labial gland biopsy. Histopathology: positive parotid or labial gland biopsy was defined as focus score ≥1.

Table 2 Ultrasound of major salivary glands versus classification criteria

	AECG	
	Parotid gland biopsy	**Labial gland biopsy**
Optimal cut-off point	15/16	15/16
% Absolute agreement	82.4% (80/97)	79.2% (76/96)
Sensitivity	71.1% (32/45)	67.3% (33/49)
Specificity	92.3% (48/52)	91.4% (43/47)
PPV	88.9% (32/36)	89.2% (33/37)
NPV	78.7% (48/61)	72.9% (43/59)
	ACR	
	Parotid gland biopsy	Labial gland biopsy
Optimal cut-off point	15	15
% Absolute agreement	85.6% (83/97)	81.7% (76/93)
Sensitivity	77.3% (34/44)	69.8% (30/43)
Specificity	92.4% (49/53)	92.0% (46/50)
PPV	89.4% (34/38)	88.2% (30/34)
NPV	83.1% (49/59)	78.0% (46/59)
	ACR-EULAR	
	Parotid gland biopsy	Labial gland biopsy
Optimal cut-off point	15	15
% Absolute agreement	79.8% (79/99)	75.3% (73/97)
Sensitivity	67.3% (35/52)	61.8% (34/55)
Specificity	93.6% (44/47)	92.9% (39/42)
PPV	92.1% (35/38)	91.9% (34/37)
NPV	72.1% (44/61)	65.0% (39/60)

In the left column parotid gland biopsies are considered an item of the AECG, ACR and ACR-EULAR classification criteria. In the right column labial gland biopsies are considered an item of the AECG, ACR and ACR-EULAR criteria.
ACR, American College of Rheumatology; AECG, American European Consensus Group; EULAR, European League Against Rheumatism; NPV, negative predictive value; PPV, positive predictive value.

In patients with negative sUS combined with absence of anti-SSA/Ro antibodies, 93% (37/40) had a negative parotid gland biopsy and 77% (17/22) had a negative labial gland biopsy.

In patients with positive sUS combined with anti-SSA/Ro antibodies, 94% (32/34) fulfilled the AECG, 97% (34/35) fulfilled the ACR and 97% (35/36) fulfilled the ACR-EULAR criteria, considering the outcome of parotid gland biopsies as an item for classification. Percentages are equal when the outcome of labial gland biopsy is considered as an item for classification (figure 4).

In patients with negative sUS combined with absence of anti-SSA/Ro antibodies, 98% (45/46) did not fulfil the AECG, 100% (45/45) did not fulfil the ACR and 98% (44/45) did not fulfil the ACR-EULAR criteria, considering the outcome of parotid gland biopsies as an item for classification (figure 4). Percentages are lower when the outcome of labial gland biopsy is considered as an item for classification: 89% (40/45) did not fulfil the AECG, 93% (43/46) did not fulfil the ACR and 89% (39/44) did not fulfil the ACR-EULAR criteria (figure 4).

DISCUSSION

This study assessed the validity of sUS in a representative population of patients with clinically suspected pSS. It is the first study that directly compared the validity of sUS with parotid gland biopsy outcome, and to the best of our knowledge the first that compared the validity of sUS with the ACR-EULAR criteria.

Since the Hocevar scoring system was the most comprehensive and valid, it was incorporated in our daily clinical practice in order to address all relevant aspects of sUS. Previously, three studies evaluated the optimal cut-off of the Hocevar scoring system.[12] These studies used different gold standards, viz AECG criteria,[12] labial gland biopsy[19] or AECG and ACR criteria.[20] The cut-off points ranged from 15 to 19. In the present study, the optimal cut-off point for the Hocevar score was found to be 15 in almost all analyses. Applying a higher cut-off value, for example, 17 as used by Hocevar *et al*,[12] would lead to increased specificity and PPV for parotid and labial gland biopsies, as well as for the AECG, ACR and ACR-EULAR classification criteria, at the expense of sensitivity and NPV (see online supplementary table S2 and S3). Furthermore, when comparing the PPV and NPV reported in different studies, it is important to keep in mind that both strongly depend on the prevalence of the disease in the investigated population, while sensitivity and specificity are less influenced.

To date, no consensus has been reached about the optimal scoring system and cut-off point for the ultrasonographic evaluation of the major salivary glands in patients suspected with pSS. Consensus is certainly needed to further elucidate the role of sUS in the diagnosis of pSS and is prerequired for being officially included in the classification criteria for pSS. Additionally, it will make the comparison between similar studies easier. sUS showed good absolute agreement and specificity using salivary gland biopsies as gold standard. Both were slightly higher when parotid gland instead of labial gland biopsies were used, possibly because the parotid gland is included in the sUS evaluation, whereas the labial gland is not.

Interestingly, the parotid gland biopsies were negative in most patients with a negative sUS, but the labial gland biopsies were positive in 26% of patients with a negative sUS. On the other hand, positive sUS predicts positive labial gland biopsies, while 22% of patients with a positive sUS had a negative parotid gland biopsy. When comparing the results of parotid and labial gland biopsies, it is important to keep in mind not all patients had both biopsies performed and the populations in which the agreement of sUS with parotid and labial gland biopsies were evaluated were not the same. However, this is the first study to compare sUS with parotid as well as labial gland biopsies.

There was a fair reversed association between total ultrasound score and UWS total flow. In other words, patients with more pronounced abnormalities on ultrasound tend to have a reduced UWS production. However, there was also a significant amount of patients with few abnormalities on ultrasound who did have a reduced UWS production. This group may consist of patients with early pSS, where UWS is lowered but no sUS abnormalities are yet seen or another condition is causing the decrease in UWS. A next step would be to assess whether the performance of the different classification criteria will improve, when UWS is replaced by salivary gland ultrasonography.

The specificity of sUS was high when anti-SSA/Ro antibody status was used as gold standard, confirming a previous study.[21]

Our findings regarding the sensitivity and specificity of sUS outcome compared with the classification criteria for pSS are similar with the results described in a recent meta-analysis.[8] Sensitivity of sUS was lowest when compared with the recently published ACR-EULAR criteria, when either parotid or labial gland biopsies were considered as an item for classification. It is currently unknown if sUS is sensitive enough to detect changes in the major salivary glands early in the disease course. In our inception cohort, more patients with low sUS scores fulfil the ACR-EULAR criteria compared with the AECG and ACR criteria. This may either suggest that patients are classified as pSS at an earlier stage of the disease according to the ACR-EULAR criteria or that the ACR-EULAR criteria are too liberal.[7] Finally,

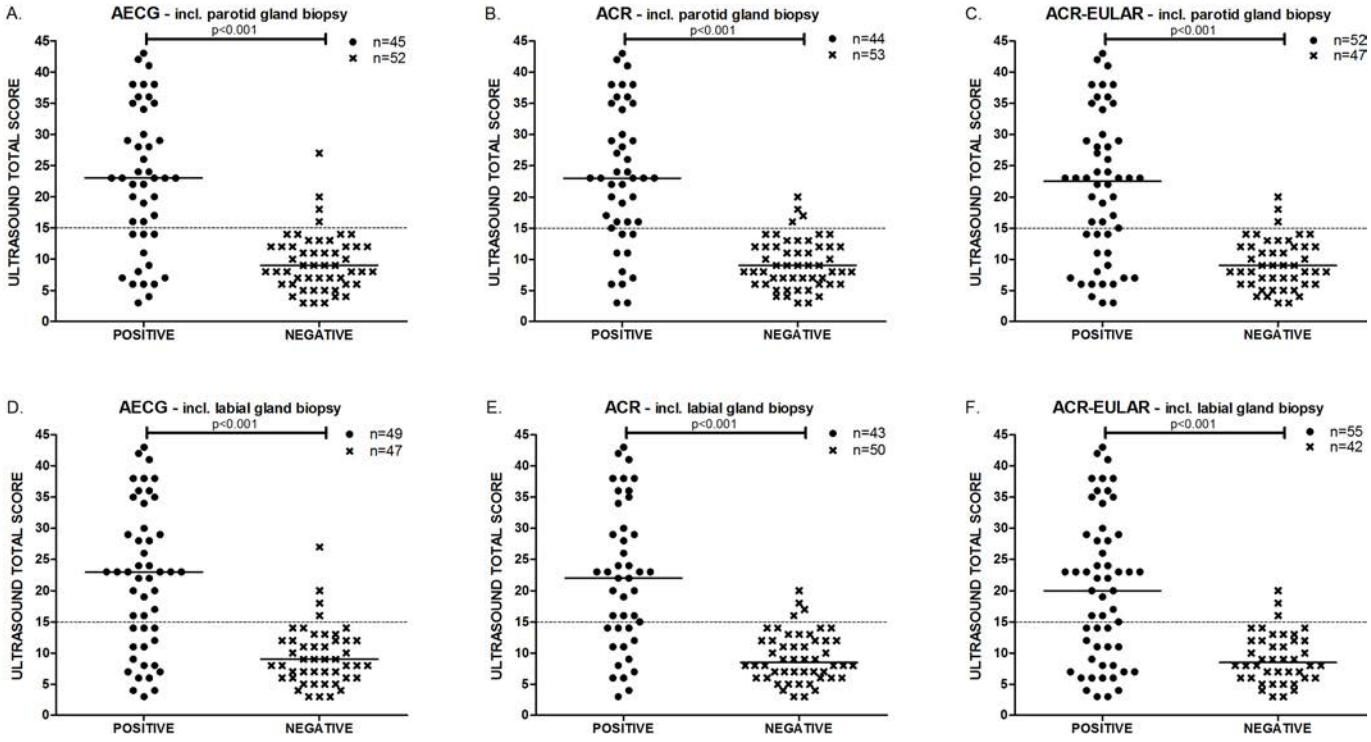


Figure 3 Ultrasound total score compared with classification criteria for pSS. In A–C, when 'including parotid gland biopsy' is added, parotid gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. In D–F, when 'including labial gland biopsy' is added, labial gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. (A) Positive versus negative of fulfilling AECG criteria, (B) positive versus negative of fulfilling ACR criteria, (C) positive versus negative of fulfilling ACR-EULAR criteria, (D) positive versus negative of fulfilling AECG criteria, (E) positive versus negative of fulfilling ACR criteria and (F) positive versus negative of fulfilling ACR-EULAR criteria. In most patients who underwent only one salivary gland biopsy, classification as pSS or non-pSS could be determined, as the biopsy would not have been decisive for classification. Therefore, ≥90% (93/103) of the patients could be included in the analysis of the agreement between sUS and the different criteria sets including parotid or labial gland biopsy. ACR, American College of Rheumatology; AECG, American European Consensus Group; EULAR, European League Against Rheumatism; pSS, primary Sjögren's syndrome; sUS, ultrasound of major salivary glands.

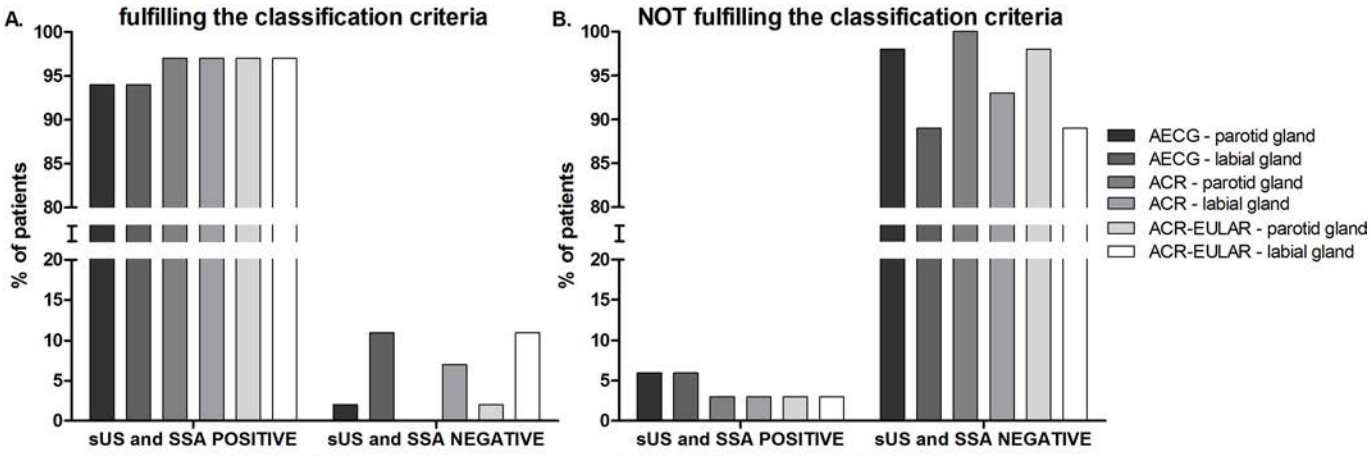
the PPV of sUS compared with the classification criteria was higher than the NPV. Thus, positive sUS predicts fulfilment of the AECG, ACR and ACR-EULAR criteria, but negative sUS does not exclude classification. When applying classification criteria, we should keep in mind that they are developed for research purposes to define homogeneous study groups, although in clinical practice they are often used for diagnostic purposes.[22]

In accordance with our findings, Astorri *et al*[23] reported that positive sUS was highly predictive of positive labial gland biopsies. Therefore, one could consider not performing a labial gland biopsy in patients with a positive sUS. This previous study also showed that negative sUS was highly predictive of negative labial gland biopsies in patients with sicca symptoms. However, we were unable to confirm this observation with our data, as we found a moderate NPV of sUS for labial gland biopsies. Therefore, in patients with a negative sUS the result of labial gland biopsies could not fully be predicted. There are some possible explanations for this discrepancy. None of the 'non-Sjögren sicca patients' in the study of Astorri *et al*[23] had a positive labial gland biopsy, while some of our 'non-Sjögren sicca patients' did have a positive biopsy. These biopsies may be false-positive,[24 25] or these patients have SS not meeting the classification criteria. Moreover, Astorri *et al*[23] did not mention the time interval between sUS and labial gland biopsy, which may be longer than in our study and a different sUS scoring system was used.

Astorri *et al*[23] stated that labial gland biopsies should not be performed in extractable nuclear antigen (ENA)-negative patients with negative sUS, unless there are other strong clinical indications for SS. Based on our data, we cannot support this conclusion. It is well established that negative serology occurs in 10%–50% of patients with pSS and correlates with milder disease.[26–28] Interestingly, when labial gland biopsies are considered an item of the classification criteria, 7%–11% of our patients with the combination of negative sUS and absence of anti-SSA/Ro antibodies were classified as pSS according to the different criteria sets. In these anti-SSA/Ro-negative patients, positive biopsies are decisive for classification.[29] Thus, ultrasound cannot fully replace salivary gland biopsies as there is a risk of underdiagnosing serological negative patients. Consequently, we recommend that physicians should still consider performing biopsies in patients with absence of anti-SSA/Ro antibodies and negative sUS, especially when gland function (eg, abnormal Schirmer or UWS) is impaired or when there are other signs and symptoms pointing to pSS.

Strikingly, almost all of our patients with both a positive sUS and presence of anti-SSA/Ro antibodies fulfilled the classification criteria for pSS. In this patient group, physicians could consider to skip the salivary gland biopsy as the combination of positive sUS and presence of anti-SSA/Ro antibodies is already highly suggestive of pSS. These rather interesting results are to be confirmed in other cohorts.

The main strength of our study is that consecutive patients clinically suspected with pSS were included. Thus, the study population clearly represents the clinical circumstances in daily clinical practice. Moreover, our Sjögren's expertise centre is one of the few centres in which both parotid and labial gland biopsies



A. fulfilling the classification criteria

	sUS/SSA pos	sUS/SSA neg
AECG - parotid gland	94% (32/34)	2% (1/46)
AECG - labial gland	94% (33/35)	11% (5/45)
ACR - parotid gland	97% (34/35)	0% (0/45)
ACR - labial gland	97% (30/31)	7% (3/46)
ACR-EULAR - parotid gland	97% (35/36)	2% (1/45)
ACR-EULAR- labial gland	97% (34/35)	11% (5/44)

B. NOT fulfilling the classification criteria

	sUS/SSA pos	sUS/SSA neg
AECG - parotid gland	6% (2/34)	98% (45/46)
AECG - labial gland	6% (2/35)	89% (40/45)
ACR - parotid gland	3% (1/35)	100% (45/45)
ACR - labial gland	3% (1/31)	93% (43/46)
ACR-EULAR - parotid gland	3% (1/36)	98% (44/45)
ACR-EULAR- labial gland	3% (1/35)	89% (39/44)

Figure 4 Percentage of patients either or not fulfilling the AECG, ACR or ACR-EULAR criteria including only parotid gland biopsy or only labial gland biopsy. When 'parotid gland' is mentioned, parotid gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. When 'labial gland' is mentioned, labial gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. (A) Fulfilling the AECG, ACR or ACR-EULAR classification criteria, with on the left positive sUS in combination with presence of anti-SSA/Ro antibodies and on the right negative sUS in combination with absence of anti-SSA/Ro antibodies. (B) NOT fulfilling the AECG, ACR or ACR-EULAR classification criteria, with on the left positive sUS in combination with presence of anti-SSA/Ro antibodies and on the right negative sUS in combination with absence of anti-SSA/Ro antibodies. ACR, American College of Rheumatology; AECG, American European Consensus Group; EULAR, European League Against Rheumatism; sUS, ultrasound of major salivary gland.

can be performed.[14] Having access to parotid gland biopsies has several advantages, in particular that repeated biopsies of the same parotid gland can be performed (eg, for monitoring treatment efficacy) and mucosa-associated lymphoid tissue (MALT)-lymphoma might be identified at an earlier stage.[14]

In conclusion, in our prospective inception cohort study derived from daily clinical practice, absolute agreement between sUS and salivary gland biopsies was good, although slightly higher for parotid compared with labial gland biopsies.

1. Positive sUS predicts classification according to the AECG, ACR and ACR-EULAR classification criteria, but negative sUS does not exclude classification.
2. Positive sUS in combination with presence of anti-SSA/Ro antibodies highly predicts classification according to the AECG, ACR and ACR-EULAR criteria. The combination of negative sUS and absence of anti-SSA/Ro antibodies highly excludes classification when parotid gland biopsy outcome is considered as an item for classification, but when the outcome of labial gland biopsy is considered as an criteria item, combining negative sUS with absence of anti-SSA/Ro antibodies does not exclude classification.

Acknowledgements The authors would like to acknowledge the participating patients. Part of this study has been presented at the ACR 2016 in Washington and the corresponding abstract has been published as a conference abstract.

Contributors EM and KD: designed the study, collected, analysed and interpreted data, drafted the article, and approved the version to be published. SA: designed the study, analysed and interpreted data, critically revised the article, and approved the version to be published. JFvN and AJS: critically revised the article and approved the version to be published. FGMK, FKLS, AV and HB: designed the study, critically revised the article and approved the version to be published.

Funding No funding has been received for the conduct and/or preparation of this article.

Competing interests None declared.

Patient consent Detail has been removed from this case description/these case descriptions to ensure anonymity. The editors and reviewers have seen the detailed information available and are satisfied that the information backs up the case the authors are making.

Ethics approval Ethics committee of the University Medical Center Groningen (METC waiver 016/120).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data available.

REFERENCES

1 Qin B, Wang J, Yang Z, *et al*. Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. *Ann Rheum Dis* 2015;74:1983–9.
2 Fox RI. Sjogren's syndrome. *Lancet* 2005;366:321–31.
3 Vitali C, Bombardieri S, Jonsson R, *et al*. Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. *Ann Rheum Dis* 2002;61:554–8.
4 Shiboski CH, Shiboski SC, Seror R, *et al*. American College of Rheumatology/European League against Rheumatism classification criteria for primary Sjögren's syndrome. *Ann Rheum Dis* 2016;2017:9–16.
5 Shiboski SC, Shiboski CH, Criswell L, *et al*. American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance cohort. *Arthritis Care Res* 2012;64:475–87.
6 Vitali C, Bootsma H, Bowman SJ, *et al*. Classification criteria for Sjogren's syndrome: we actually need to definitively resolve the long debate on the issue. *Ann Rheum Dis* 2013;72:476–8.
7 Vissink A, Bootsma H. Refining the classification criteria for primary Sjögren's syndrome. *Nat Rev Rheumatol* 2017;13:10–12.
8 Delli K, Dijkstra PU, Stel AJ, *et al*. Diagnostic properties of ultrasound of Major salivary glands in Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792–800.
9 Jousse-Joulin S, Milic V, Jonsson MV, *et al*. Is salivary gland ultrasonography a useful tool in Sjögren's syndrome? A systematic review. *Rheumatology* 2016;55:789–800.

10 Cornec D, Jousse-Joulin S, Marhadour T, *et al*. Salivary gland ultrasonography improves the diagnostic performance of the 2012 American College of Rheumatology classification criteria for Sjögren's syndrome. *Rheumatology* 2014;53:1604–7.
11 Theander E, Mandl T. Primary Sjögren's syndrome: diagnostic and prognostic value of salivary gland ultrasonography using a simplified scoring system. *Arthritis Care Res* 2014;66:1102–7.
12 Hocevar A, Ambrozic A, Rozman B, *et al*. Ultrasonographic changes of major salivary glands in primary Sjogren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology* 2005;44:768–72.
13 Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. *J Clin Pathol* 1968;21:656–60.
14 Pijpe J, Kalk WW, van der Wal JE, *et al*. Parotid gland biopsy compared with labial biopsy in the diagnosis of patients with primary Sjogren's syndrome. *Rheumatology* 2007;46:335–41.
15 Greenspan JS, Daniels TE, Talal N, *et al*. The histopathology of Sjögren's syndrome in labial salivary gland biopsies. *Oral Surg Oral Med Oral Pathol* 1974;37:217–29.
16 Jensen SB, Vissink A. Salivary gland dysfunction and xerostomia in Sjögren's syndrome. *Oral Maxillofac Surg Clin North Am* 2014;26:35–53.
17 Hosmer DW, Lemeshow S, Sturdivant RX, *et al*. *Applied Logistic Regression*. 2013. Third.
18 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
19 Milic VD, Petrovic RR, Boricic IV, *et al*. Diagnostic value of salivary gland ultrasonographic scoring system in primary Sjogren's syndrome: a comparison with scintigraphy and biopsy. *J Rheumatol* 2009;36:1495–500.
20 Zhang X, Zhang S, He J, *et al*. Ultrasonographic evaluation of Major salivary glands in primary Sjögren's syndrome: comparison of two scoring systems. *Rheumatology* 2015;54:1680–7.
21 Luciano N, Baldini C, Tarantini G, *et al*. Ultrasonography of Major salivary glands: a highly specific tool for distinguishing primary Sjögren's syndrome from undifferentiated connective tissue diseases. *Rheumatology* 2015;54:2198–204.
22 Bootsma H, Spijkervet FK, Kroese FG, *et al*. Toward new classification criteria for Sjögren's syndrome? *Arthritis Rheum* 2013;65:21–3.
23 Astorri E, Sutcliffe N, Richards PS, *et al*. Ultrasound of the salivary glands is a strong predictor of labial gland biopsy histopathology in patients with sicca symptoms. *J Oral Pathol Med* 2016;45:450–4.
24 Radfar L, Kleiner DE, Fox PC, *et al*. Prevalence and clinical significance of lymphocytic foci in minor salivary glands of healthy volunteers. *Arthritis Rheum* 2002;47:520–4.
25 Segerberg-Konttinen M, Konttinen YT, Bergroth V. Focus score in the diagnosis of Sjögren's syndrome. *Scand J Rheumatol Suppl* 1986;61:47–51.
26 Peri Y, Agmon-Levin N, Theodor E, *et al*. Sjögren's syndrome, the old and the new. *Best Pract Res Clin Rheumatol* 2012;26:105–17.
27 Ramos-Casals M, Solans R, Rosas J, *et al*. Primary Sjögren syndrome in Spain. *Medicine* 2008;87:210–9.
28 Oni C, Mitchell S, James K, *et al*. Eligibility for clinical trials in primary Sjögren's syndrome: lessons from the UK Primary Sjögren's Syndrome Registry. *Rheumatology* 2016;55:544–52.
29 Barone F, Campos J, Bowman S, *et al*. The value of histopathological examination of salivary gland biopsies in diagnosis, prognosis and treatment of Sjögren's Syndrome. *Swiss Med Wkly* 2015;145.

OPEN ACCESS

EXTENDED REPORT

Minimal to no transfer of certolizumab pegol into breast milk: results from CRADLE, a prospective, postmarketing, multicentre, pharmacokinetic study

Megan EB Clowse,[1] Frauke Förger,[2] Caroline Hwang,[3] John Thorp,[4] Radboud JEM Dolhain,[5] Astrid van Tubergen,[6] Laura Shaughnessy,[7] Jeff Simpson,[7] Marie Teil,[8] Nathalie Toublanc,[9] Maggie Wang,[7] Thomas W Hale[10]

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-2017-211384).

[1]Duke University School of Medicine, Durham, North Carolina, USA
[2]Department of Rheumatology and Immunology and Allergology, Inselspital, University Hospital and University of Bern, Bern, Switzerland
[3]Keck Hospital of USC, Los Angeles, California, USA
[4]University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[5]University Medical Centre Rotterdam, Rotterdam, The Netherlands
[6]Maastricht University Medical Center, Maastricht, The Netherlands
[7]UCB Pharma, Raleigh, North Carolina, USA
[8]UCB Pharma, Slough, UK
[9]UCB Pharma SA, Braine, Belgium
[10]Texas Tech University School of Medicine, Amarillo, Texas, USA

Correspondence to
Dr Megan EB Clowse, Duke University School of Medicine, Trent Drive Durham, North Carolina 27710, USA; megan.clowse@dm.duke.edu

Received 25 February 2017
Revised 22 May 2017
Accepted 1 July 2017
Published Online First 15 August 2017

To cite: Clowse MEB, Förger F, Hwang C, *et al*. *Ann Rheum Dis* 2017;**76**:1890–1896.

ABSTRACT

Background Women with chronic inflammatory diseases face uncertainty regarding the safety of biologics during breast feeding. CRADLE was the first industry-sponsored study to evaluate certolizumab pegol (CZP) concentrations in human breast milk and estimate average daily infant dose (ADID) of maternal CZP.

Methods CRADLE (NCT02154425) was a pharmacokinetic study of lactating mothers receiving CZP. After ≥3 CZP doses, breast milk samples were collected across one dosing period (14 days for 200 mg every 2 weeks [Q2W]; 28 days for 400 mg every 4 weeks [Q4W]). Optimal analytical methods were developed to determine CZP and polyethylene glycol (PEG) levels in breast milk. ADID and relative infant dose (RID) were estimated. Safety events in mothers and infants were assessed.

Results 19 CZP-treated mothers were screened; 17 entered the sampling period: 16 on 200 mg Q2W, 1 on 400 mg Q4W. 77/137 (56%) breast milk samples had no measurable CZP. For 4/17 mothers, all samples were below the lower limit of quantification (LLOQ). Estimated ADID was 0–0.0104 mg/kg/day; median RID: 0.15%. PEG was undetectable in 134/137 samples (results could not be determined in three samples). Infants of CZP-exposed mothers had a safety profile consistent with that of unexposed similar-age infants.

Conclusion When quantifiable, CZP concentrations were <3× LLOQ (<1% plasma concentration observed with therapeutic dose), indicating no/minimal CZP transfer from plasma to breast milk. RID was 0.15% of maternal dose; <10% is considered unlikely to be of clinical concern. No PEG transfer was observed. CZP absorption by infants via breast milk is unlikely due to its low oral bioavailability and Fc-free molecular structure. These findings are reassuring and support continuation of CZP treatment during breast feeding.

Trial registration number NCT02154425; Results.

INTRODUCTION

Women with chronic inflammatory diseases, such as rheumatoid arthritis (RA), ankylosing spondylitis (AS), axial spondyloarthritis (axSpA), psoriatic arthritis (PsA), and Crohn's disease (CD) often experience postpartum flares.[1–4] Treatment of these conditions frequently involves monoclonal antibodies, such as anti-tumour necrosis factor (anti-TNF), and diffusion of these molecules into breast milk has been reported.[5 6] Existing evidence on anti-TNF transfer into breast milk lacks robust and systematic sample collection and assays validated in breast milk; reports are restricted to a few studies with a limited number of patients receiving infliximab,[7] adalimumab,[7 8] etanercept[9] or certolizumab pegol (CZP).[10 11] The lack of systematic collection of breast milk samples throughout dosing intervals, coupled with the absence of drug-specific assays validated in breast milk, suggests that existing data cannot be confidently translated into evidence-based clinical practice. Consequently, women treated with monoclonal antibodies who are considering breast feeding, as well as their physicians, face uncertainty regarding drug safety.[12]

Breast feeding is extremely important to child health and development.[13 14] Immunological and anti-inflammatory agents are passed on to the infant via breast milk, allowing development of protective mechanisms against several diseases.[15] In addition to creating an emotional bond between mother and child at an early stage, breast feeding has been associated with a decreased risk for sudden infant death syndrome[16] and other conditions.[17] Despite these benefits, the conflict between the risks of maternal medications needed for postpartum disease flare and ensuring optimal child nutrition through breast feeding presents a complex challenge.

Although biologics generally have very low oral bioavailability due to their large molecular size and the proteolytic environment in the digestive system,[18] the neonatal Fc receptor on human intestinal epithelial cells may promote uptake of undigested immunoglobulins (IgGs). CZP, the only PEGylated anti-TNF without an Fc region, has demonstrated efficacy for the treatment of RA,[19] CD,[20] axSpA[21] and PsA.[22] Physiologically, only minimal amounts of CZP are likely to cross into breast milk and be absorbed by the infant, due to its large molecule size and the replacement of the Fc portion with polyethylene glycol (PEG).[23]

The primary aim of CRADLE (NCT02154425), the first industry-sponsored, multicentre study to evaluate transfer of a biologic into breast milk, was to determine the concentrations of CZP in mature breast milk and to calculate the average daily infant dose (ADID), which is the daily CZP dose potentially ingested by the infant. Including multiple predefined sampling time points throughout the dosing interval allowed full characterisation of the

CZP pharmacokinetics (PK) in mature milk at steady state. The exploratory aim was to determine the breast milk concentrations of PEG. The relative infant dose (RID), which estimates the theoretical infant dose as a percentage of the weight-normalised maternal dose was calculated post hoc. Safety events in mothers and infants were examined.

METHODS

Study design and patients

CRADLE (NCT02154425) was a prospective, postmarketing, multicentre PK study to measure the CZP concentration in breast milk. It enrolled lactating mothers, at least 6 weeks postpartum with no upper age limit for infants, receiving commercial CZP for an approved indication (RA, CD, AS/axSpA and PsA), as prescribed by their treating physician. Importantly, the decision to treat with CZP and to breastfeed were made prior to and independently from study participation. CZP was not provided by the study sponsor.

No exclusions were made regarding multiple births, but women who were pregnant, or planned to become pregnant during study duration, were ineligible to participate. Mothers with positive or indeterminate tuberculosis (TB) testing, active or latent TB infection or at high risk of TB infection were excluded, as were mothers who had received treatment with any biologic or anti-TNF other than CZP within five half-lives prior to collection of the first milk sample. Also excluded were mothers of premature infants (<37 weeks gestation). Mothers were withdrawn if they took any biological disease-modifying drug other than CZP during the sampling period. In addition, mothers with active mastitis were excluded from the sampling period until resolution.

The study was designed as a 'milk-only' study in accordance with the 2005 US Food and Drug Administration (FDA) draft guidance[24] and was conducted between September 2014 and October 2015 by six investigators: three in the USA, one in Switzerland and two in the Netherlands. In addition to traditional site-based enrolment, an open enrolment model was used in the USA and Canada, with remote oversight by a participating investigator. This enabled remotely located mothers to take part in the study without the burden of travel, thus allowing the study to reach more mothers meeting enrolment criteria than through traditional site-based enrolment alone. The study design and protocol were approved by the local ethics committee or institutional review board (IRB) at each participating site, or by a central IRB, as applicable. All mothers provided written informed consent to participate.

Study procedures

Mothers received commercial CZP on either the 2-weekly dose (CZP 200 mg every 2 weeks [Q2W]) or the 4-weekly dose regimen (CZP 400 mg every 4 weeks [Q4W]). After at least three CZP doses, when plasma CZP concentration in the mothers was considered at steady-state, mature breast milk samples were collected across a single dosing period. For mothers on CZP 200 mg Q2W, samples were collected on days 0, 2, 4, 6, 8, 10, 12 and 14; for mothers on CZP 400 mg Q4W, an additional sample was collected on day 28 (figure 1). All mothers were to dose on the same day as the first sample collection, with date and time of dosing being recorded by the home healthcare nurse. Predose samples were those collected on days 0 and 14 for mothers on CZP 200 mg Q2W, and on days 0 and 28 for mothers on CZP 400 mg Q4W. Compliance with CZP therapy was not mandated in the study; however, at day 0, all mothers were confirmed to have received at least three doses of CZP prior to the beginning of the sampling period.

In-home nursing visits for milk sample collection minimised the burden on mothers and were carried out at all study sites, using the same procedure: at each visit and at approximately the same time of day, milk was pumped from both breasts until completely emptied (accommodating the mothers' schedule), using the electronic breast pump (Medela Advanced Personal Double Pump) provided by the sponsor. Milk from both breasts was mixed by the home healthcare nurse, before 5 mL were placed on dry ice and shipped to the central laboratory (Quintiles; Atlanta [USA sites] or Edinburgh, UK [European sites]). During each visit, the nurse confirmed whether the infant had received any other nutrition in the prior 2 days to explore whether there was any difference between women who breast fed exclusively and those that supplemented breast feeding with other nutrition sources (eg, formula milk). To avoid interference by PEG-containing creams, mothers were only permitted to use PEG-free nipple/breast creams provided by the sponsor.

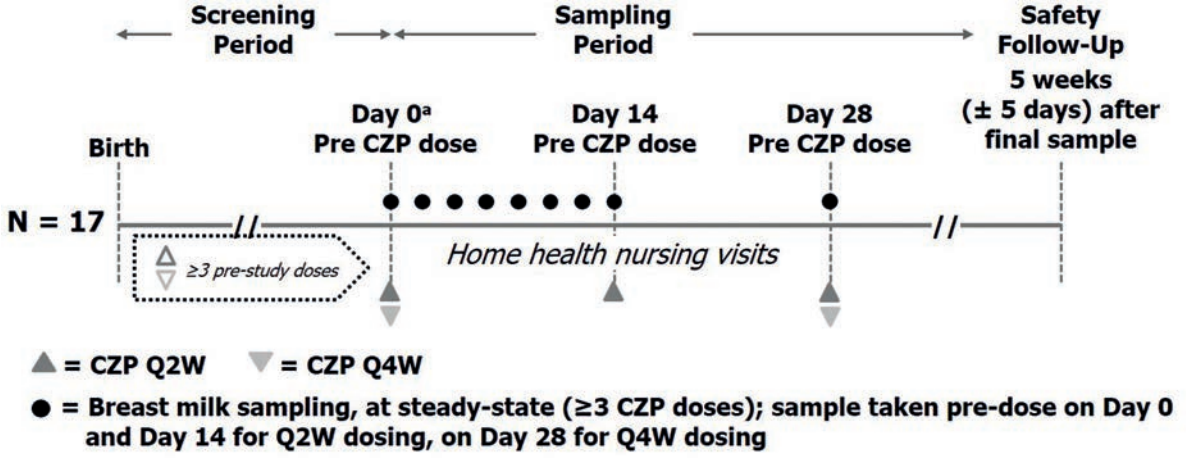


Figure 1 CRADLE study design. CZP, certolizumab pegol; Q2W, every 2 weeks; Q4W, every 4 weeks.

Samples were subsequently analysed at Covance Inc. (Chantilly, Virginia, USA). CZP concentration in breast milk was measured using an electro-chemiluminescence immunoassay, in which CZP was captured by a TNF-coated multiarray electrode and detected with an anti-PEG antibody, prior to reading on a Meso Scale Discovery (MSD; Rockville, Maryland, USA) platform.[25] The assay is CZP-specific and, due to the technical advantages of the MSD methodology (high sensitivity, large dynamic range, small sample volume[26 27]), >10 times more sensitive (lower limit of quantification [LLOQ]: 0.032 µg/mL) than the previous ELISA used in other CZP PK studies.[11 28 29] The assay was validated in milk; CZP stability in milk was confirmed.

The concentrations of total PEG (ie, PEG present as intact CZP or in deconjugated form) were determined by a validated assay using nuclear magnetic resonance spectroscopy (LLOQ: 0.5 µg/mL).[30]

Study endpoints

The primary objectives were to determine CZP concentrations in breast milk and to calculate the ADID of maternal CZP. The exploratory objective was to assess the concentration of total PEG in breast milk. A post hoc variable, RID, was calculated.[31] RID is the infant dose as a per cent of the weight-related maternal dose and is widely used by lactation specialists, paediatricians and neonatologists to assess risk to infants.[31 32] Analysis of the PK parameters was performed using Phoenix WinNonlin V.6.4 (Certara, USA).

The safety analysis included adverse events (AEs) in all mothers who received at least one dose of CZP and the infants of all participating mothers from the time of informed consent through safety follow-up (up to 5 weeks±5 days after the final sample was obtained). Patient consent was obtained after delivery, up to 10 weeks before day 0. Prior to sampling on day 0, eligibility was reconfirmed. During each sampling visit, patients were given the opportunity to report AEs spontaneously, and a general prompt using open-ended questions was also given. If an AE was reported during the in-home visit, the home healthcare nurse contacted the principal investigator to speak directly with the patient for further assessment.

AEs of interest included any opportunistic infections, malignancies (including unspecified), major adverse cardiac events, haematopoietic cytopaenias, serious bleeding, hepatic events and injection reactions (local or systemic).[32] AEs were coded using MedDRA V.18.1.

Statistical analysis

No formal sample size calculations were performed, as no statistical hypotheses were being tested. The planned sample size for the study was 16 mothers; approximately twice the minimum number of subjects considered sufficient by the FDA for a 'milk-only' study.[24] Summary statistics were reported for quantitative variables and frequency tables for qualitative data. Statistical analysis was performed using SAS V.9.3. All summaries of PK variables were based on the values observed at each visit: no imputation was used.

In addition to CZP concentrations, three measures of CZP transfer were considered: (1) the average daily CZP concentration in breast milk (C_{ave}), (2) ADID and (3) RID. C_{ave} is calculated by non-compartmental analysis from the concentration versus time profile over the dosing interval, using the actual sampling days. ADID is the dose of CZP ingested by a child based on C_{ave} and the estimated daily volume of milk ingested. As per the 2005 FDA draft guidance,[24] the standardised milk consumption of a fully breastfed 2-month-old infant (150 mL/kg/day) was used to calculate the ADID over a dosing interval (14 or 28 days) from the study data:

$$Estimated\ ADID\ \left(mg/kg/day\right) = C_{average} \times 150\,mL/kg/day$$

The exploratory post hoc variable, RID, was calculated as follows[31]:

$$RID\ (\%) = \frac{ADID\ \left(mg/kg/day\right)}{Maternal\ dose\ \left(mg/kg/day\right)} \times 100$$

RID was not calculated when all results were below the lower limit of quantification (BLQ), as was the case in four mothers.

Subgroup analyses were performed, based on the mothers' indications (RA, CD and PsA) and on the use of supplemental nutrition (non-exclusive vs exclusive breast feeding) for the PK parameters. Subgroup analyses were not performed on groups with fewer than three patients (axSpA).

RESULTS

Patients

Between 8 September 2014 and 30 October 2015, 19 mothers were screened; 18 received commercial CZP and met the inclusion criteria (one mother was prescribed CZP prior to screening, but did not receive CZP and was therefore ineligible). One mother failed screening and discontinued the study due to an AE of herpes zoster. All 17 mothers who entered the sampling period completed the study (no missed visits);

Table 1 Demographics and baseline characteristics of mothers and infants

	All mothers (n=18)*
Mean (SD), unless otherwise stated	
Age, years	33.7 (4.2)
Weight, kg	68.9 (9.6)†
BMI, kg/m²	23.6 (3.0)†
Location, n	
USA/Canada‡	10
Switzerland	5
The Netherlands	3
Mother's indication for CZP treatment, n†	
Rheumatoid arthritis	7
Crohn's disease	5
Psoriatic arthritis	3
Axial spondyloarthritis/ankylosing spondylitis	2
	All infants (n=17)
Median (min–max), unless otherwise stated	
Female, n (%)	11 (64.7)
Gestational age at birth, weeks	40.0 (39.0–41.7)
Weight at birth, kg	3.5 (2.6–4.1)
Length at birth, cm	50.7 (48.0–57.0)
Age at time of mother's first sample, months	2.8 (1.6–16.8)
Age at time of mother's first sample, n (%)	
≤6 months	13 (76.5)
>6 months–≤12 months	2 (11.8)
≥12 months–≤18 months	2 (11.8)

*Includes one screen failure.
†n=17.
‡One Canadian patient enrolled under the central USA site, which was approved by the Canadian central IRB.
BMI, body mass index; IRB, institutional review board.

Table 2 Concentrations of CZP (µg/mL) in breast milk after administration of CZP dose in mothers

Mother number	Relative time (days)								
	0	2	4	6	8	10	12	14	28
17	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ
4	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	–
13	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	–
14	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	BLQ	–
7	BLQ	BLQ	BLQ	BLQ	BLQ	0.035	BLQ	BLQ	–
6	BLQ	BLQ	0.044	0.048	BLQ	BLQ	BLQ	BLQ	–
8	BLQ	BLQ	0.035	0.034	0.043	BLQ	BLQ	BLQ	–
10	BLQ	BLQ	BLQ	0.033	0.042	0.042	BLQ	BLQ	–
12	BLQ	BLQ	0.034	0.037	0.033	BLQ	BLQ	BLQ	–
2	BLQ	BLQ	0.035	0.037	0.041	BLQ	0.043	BLQ	–
11	BLQ	BLQ	0.051	0.038	0.042	BLQ	0.033	BLQ	–
15	BLQ	BLQ	0.041	0.034	0.033	BLQ	0.037	BLQ	–
16	0.040	0.033	0.036	0.037	0.043	BLQ	BLQ	BLQ	–
3	BLQ	0.032	0.049	0.053	0.037	0.037	0.033	0.033	–
9	0.039	0.040	0.047	0.045	0.042	0.043	0.038	0.035	–
1	0.057	0.051	0.066	0.065	0.062	0.056	0.052	0.041	–
5	0.056	0.069	0.074	0.076	0.076	0.069	0.069	0.060	–

White coloured areas depict CZP concentration Less than 3×LLOQ (<0.096 µg/mL).
Grey coloured areas depict CZP concentration Less than 2×LLOQ (<0.064 µg/mL).
Dark grey coloured areas depict CZP concentration BLQ (<0.032 µg/mL).
Days 0 and 14 are predose for mothers on the CZP 200 mg Q2W dosing regimen.
Days 0 and 28 are predose for the mother on the CZP 400 mg Q4W dosing regimen.
For reference, the mean 12-week CZP plasma C_{trough} value, that is, the trough concentration at steady-state, reported from non-pregnant patients with RA receiving CZP 200 mg Q2W in the RAPID2 trial was 15.7 µg/mL (95% CI 14.0 to 17.7).[28]
BLQ, below the lower limit of quantification, <0.032 µg/mL; CZP, certolizumab pegol; LLOQ, lower limit of quantification; Q2W, every 2 weeks; Q4W, every 4 weeks; RA, rheumatoid arthritis.

100% of planned samples were collected. Seven mothers (38.9%) were diagnosed with RA, 5 (27.8%) with CD (two of which also had rheumatic diseases), 3 (16.7%) with PsA and 2 (11.1%) with axSpA. The demographics and baseline characteristics of all mothers screened and of all infants of participating mothers are shown in table 1. The mothers' indications for CZP treatment and their infants' age at the time of first sampling are listed in online supplementary table 1.

Pharmacokinetic results

One hundred thirty-seven breast milk samples were collected from 17 mothers (16 on CZP 200 mg Q2W, one on CZP 400 mg Q4W). All samples had CZP concentrations that were minimal or BLQ. Importantly, 77/137 (56%) samples had no measurable CZP (table 2). Four of the 17 women, including one on CZP 400 mg Q4W, did not have measurable CZP levels in their breast milk at any time point. In the remaining 13 mothers with a measurable concentration during at least one time point, the highest CZP concentration was 0.076 µg/mL, less than three times LLOQ (0.032 µg/mL).

The estimated ADID ranged from 0 to 0.0104 mg/kg/day; median estimated ADID was 0.003503 mg/kg/day (table 3). Additional CZP PK parameters are shown in table 3. The RID, calculated post hoc,[31 32] ranged from 0.04% to 0.30%; median RID was 0.15%.

PK parameters were similar for women with different indications as well as between mothers using supplemental nutrition and those breast feeding exclusively.

Table 3 Pharmacokinetic parameters of CZP in breast milk after administration of CZP

Parameter	n	Geo. mean (Geo. CV (%))	Median (min–max)
All mothers: both CZP 200 mg Q2W and CZP 400 mg Q4W dosing regimens			
Estimated average daily infant dose (mg/kg/day)	17	0.00426 (59.4)	0.003503 (0–0.0104)
C_{ave} (µg/mL)	17	0.0248 (58.0)	0.02335 (0.00744–0.0693)
t_{max} (day)	13	–	5.051 (2.89–11.9)
Mothers on CZP 200 mg Q2W dosing regimen only			
AUC_T (day* µg/mL)	13	0.398 (59.4)	0.4249 (0.104–0.970)
C_{max} (µg/mL)	16	0.0383 (50.3)	0.04285 (BLQ–0.0758)

AUC_T, area under the curve over a dosing interval (14 or 28 days); C_{ave}, average concentration over a dosing interval; C_{max}, maximum observed CZP concentration in milk over the dosing interval; CV, coefficient of variations; CZP, certolizumab pegol; Geo.: geometric; Q2W, every 2 weeks; Q4W, every 4 weeks; t_{max}, time of the maximum observed concentration.

Table 4 Adverse events (AEs) occurring in mother-infant pairs

Mothers			Infants		
Mother (n=18)*	No. of AEs in mother (n)	AE	Infant (n=17)*	No. of AEs in infant (n)	AE
1	None		1	2	Lichen striatus
					Upper respiratory tract infection
2	1	**Breast abscess‡**	2	None	
3	None		3	1	Gastro-oesophageal reflux disease
5	1	Rash	5	1	Nasopharyngitis
7	1	Upper respiratory tract infection	7	1	Upper respiratory tract infection
8	2	Candida infection	8	1	Candida infection
		Crohn's disease flare			
10	1	Viral upper respiratory tract infection	10	None	
11	1	Headache	11	None	
13	1	Psoriatic arthritis flare	13	None	
14	2	Nipple disorder	14	2	Vomiting
		Headache			Nasopharyngitis
15	*None*		15	1	Nasopharyngitis
16	2	Upper respiratory tract infection	16	2	Upper respiratory tract infection
		Pneumonia			Nasopharyngitis
SF†	2	Herpes zoster	N/A		
		Galactostasis			
Total number of mothers experiencing any AE		10	Total number of infants experiencing any AE		8
Total number of AEs		14	Total number of AEs		11

The safety analysis included all mothers who received at least one dose of CZP and the infants of all mothers who participated in the study. The safety follow-up period extended up to 5 weeks (±5 days) after the final sample was collected. AEs in mother-infant pairs were not necessarily associated temporally.
Bold text indicates serious adverse event (SAE).
*Mother-infant pairs are numbered as per table 2.
†SF, screen failure: mother discontinued from study due to AE of herpes zoster during screening period.
‡Breast abscess during screening period, which resolved prior to sampling.
CZP, certolizumab pegol; N/A, not applicable as the mother did not enter the sampling period.

PEG concentration was analysed in 137 samples; 134 samples had no quantifiable PEG; three samples were classed as not reportable (online supplementary table 2).

Safety

The safety analysis included 18 CZP-exposed mothers and 17 infants. AEs are shown in table 4. One mother discontinued the study during the screening period due to an AE of herpes zoster. Overall, 10 mothers (55.6%) experienced 14 AEs, and 8 infants (47.1%) experienced 11 AEs.

AEs in mothers and infants were mostly mild to moderate in intensity (mothers: three mild [16.7%] and six moderate [33.3%]; infants: six mild [35.3%] and two moderate [11.8%]). One severe AE (5.6%) was reported in one mother: a breast abscess, which occurred during the screening period and was resolved prior to sampling. Five AEs in four mothers were classified as drug-related: two (11.1%) upper respiratory tract infections, and one each (5.6%) of herpes zoster, CD flare and pneumonia. Nasopharyngitis in one infant (5.9%) was considered mild in intensity and classified as drug-related by the principal investigator, based on the known CZP safety profile.[33] No serious AEs were reported in the infants. Overall, events in the mothers were consistent with the known CZP safety profile, and the events observed in infants were consistent with those in unexposed infants of similar age.[34–36]

DISCUSSION

CRADLE was the first industry-sponsored clinical lactation study evaluating the transfer of a biologic into mature breast milk of women with chronic inflammatory diseases. Because low levels of IgGs have been shown to diffuse into breast milk,[5] it is important to measure the relative abundance of therapeutic antibodies in breast milk.[32] This study found minimal transfer of CZP into breast milk, with an infant receiving, on average, 0.15% (RID) of the maternal dose.

While prior studies entailed case reports with few milk samples obtained at varying times after dosing,[37 38] CRADLE was designed to fully characterise the CZP PK profile in mature milk at steady-state. Following the FDA recommendation and in line with the 2005 guidance for industry,[24] the study was designed as a 'milk-only' study. The rationale was that the CZP PK profile was already well established[28] and that lactation was not expected to substantially change its PK. In addition, this 'milk-only' study design avoided additional burden on mothers and their babies: milk sampling did not occur earlier than 4–6 weeks postpartum to ensure that lactation and feeding patterns were well established, mature milk was being produced and maternal physiology had largely returned to pre-pregnancy state.

CZP concentration was BLQ in 56% of milk samples. When measurable, CZP concentrations were less than 3× LLOQ, with a marginal maximum concentration (0.0758 µg/mL). This is equivalent to less than 1% of the expected mean CZP plasma

trough concentration for a CZP-treated adult,[28 39] indicating no to minimal transfer of CZP from plasma to breast milk. No transfer of total PEG from plasma to breast milk was observed.

The estimated daily dose of CZP ingested by breastfed infants over the dosing interval was minimal, with maximum ADID of 0.0104 mg/kg/day. The median CZP RID was 0.15%. The number of medications available to breast feeding mothers requiring drug therapy is increasing, and RID is a useful parameter for assessing drug safety in breastfeeding by providing a standardised means of referencing infant to maternal exposure on a dose/weight basis.[31] An RID <10% is considered safe by lactation specialists, with the estimate evaluated against the potential toxicity of the drug.[31 32 40] Examples of drugs considered safe when breast feeding include analgesics (ibuprofen and acetaminophen), antibiotics (penicillins and cephalosporins), antidepressants (citalopram and sertraline) and anticoagulants (heparin and dalteparin) (RID <10%).[40]

Subgroup analyses suggest that transfer of CZP into breast milk is independent of the mother's indication and CZP dosing regimen. Additionally, no difference between subgroups was observed regarding CZP transfer into milk of exclusively breastfeeding mothers versus those who supplemented nutrition.

Although there were no enrolment restrictions for patient numbers for the two dosing regimens, 16/17 patients were on the single dose regimen, as per physician discretion. To our knowledge, this does not impact the results' relevance.

Enrolment for studies of this nature is challenging, given the unique barriers to participation, such as ethical, legal and medical considerations associated with this sensitive patient population. The CRADLE open enrolment, combined with the traditional site-based model, and including in-home nursing services, allowed mothers to participate without the burden of travel. This operational model was essential to the successful completion of the study.

We acknowledge that no preterm babies were included in this study, although it is well known that prematurity is an underlying risk for women with inflammatory diseases, particularly those with high disease activity.[41] It is noteworthy that while valuable data might be obtained from studying drug absorption in premature babies, due to the possible differences in their digestive tracts from full-term infants,[42] such an analysis would have been outside the scope of this study.

Similarly, milk collection from mothers immediately after birth was not included for a number of reasons: due to the very limited volume of breast milk/colostrum available at this time, obtaining these samples would be challenging and, in addition, the study did not wish to disrupt the initial mother/infant bonding. These limitations raise awareness for the need of further research into drug transfer from mother to infant.

No new CZP safety issues were identified. AEs in mothers were consistent with the current CZP safety profile, while events in infants were consistent with those occurring in unexposed infants of similar age. For reference, incidence for minor infections in infants is 6–8 times per year for upper respiratory tract infection/nasopharyingitis[34 35] and 2%–5% for oral candidiasis.[36]

In conclusion, these findings suggest that the level of CZP ingested by the suckling infant is minimal and indicate that continuation of CZP treatment is compatible with breast feeding. The robust clinical evidence from the CRADLE study allows breastfeeding women with chronic inflammatory diseases and their treating physicians to make informed decisions regarding their postpartum treatment.

Acknowledgements We are indebted to the mothers, and their infants, for their altruistic participation in this study. We thank the nurses and the investigator teams, the study coordination team including Samia Elyakhlifi and Lien Heinzen, UCB Pharma, and Nicole Hurst, PPD. We also acknowledge Amanda Golembesky, PhD, UCB Pharma, Raleigh, USA, and Gerry Parker, PhD, UCB Pharma, Slough, UK, for their contributions; 'Matladi N. Ndlovu, MSc, PhD, UCB Pharma, Brussels, Belgium, for publication critical review coordination; and Sana Eljamel, MBChB, Julia Bárdos, PhD, and Danalyn Byng from Costello Medical Consulting, Cambridge, UK, for medical writing and editorial assistance in preparing this manuscript for publication, based on the authors' input and direction.

Contributors Substantial contributions to study conception/design or acquisition/analysis/interpretation of data: MEBC, FF, CH, JT, RJEMD, AvT, LS, JS, MT, NT, MW and TWH; drafting of the publication or revising it critically for important intellectual content: MEBC, FF, CH, JT, RJEMD, AvT, LS, JS, MT, NT, MW and TWH; final approval of the publication: MEBC, FF, CH, JT, RJEMD, AvT, LS, JS, MT, NT, MW and TWH.

Funding UCB Pharma funded this study and manuscript. UCB Pharma reviewed only for scientific and legal accuracy.

Competing interests MEBC has received research support from Janssen and Pfizer and is a consultant for UCB Pharma; FF has received research grants from UCB Pharma and is a consultant for and/or has received speakers' fees from Mepha, Roche and UCB Pharma; CH has received research support from AbbVie and UCB Pharma and is a consultant for Janssen; JT has received research grants from UCB Pharma and is a consultant for Genentech/Roche and UCB Pharma; RJEMD has received unrestricted research grants from UCB Pharma; AvT has received research grants and/or clinical trial support from AbbVie, Celgene, Janssen Cilag, MSD, Novartis, Pfizer and UCB Pharma, has received speakers' bureau from Janssen Cilag, MSD and Pfizer and is a consultant for and/or has served on advisory boards for AbbVie, Janssen Cilag, Novartis and Pfizer; LS, JS, MT, NT and MW are employees of UCB Pharma; TWH is a consultant for UCB Pharma.

Patient consent This article does not contain personal medical information about any identifiable living individual.

Ethics approval The study design and protocol were approved by the local ethics committee or institutional review board (IRB) at each participating site, or by a central IRB, as applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 de Man YA, Dolhain RJ, van de Geijn FE, *et al*. Disease activity of rheumatoid arthritis during pregnancy: results from a nationwide prospective study. *Arthritis Rheum* 2008;59:1241–8.
2 Kurizky PS, Ferreira CC, Nogueira LS, *et al*. Treatment of psoriasis and psoriatic arthritis during pregnancy and breastfeeding. *An Bras Dermatol* 2015;90:367–75.
3 Ostensen M, Husby G. A prospective clinical study of the effect of pregnancy on rheumatoid arthritis and ankylosing spondylitis. *Arthritis Rheum* 1983;26:1155–9.
4 Kane S, Lemieux N. The role of breastfeeding in postpartum disease activity in women with inflammatory bowel disease. *Am J Gastroenterol* 2005;100:102–5.
5 Jatsyk GV, Kuvaeva IB, Gribakin SG. Immunological protection of the neonatal gastrointestinal tract: the importance of breast feeding. *Acta Paediatr Scand* 1985;74:246–9.
6 Baker TE, Cooper SD, Kessler L, *et al*. Transfer of natalizumab into breast milk in a mother with multiple sclerosis. *J Hum Lact* 2015;31:233–6.
7 Fritzsche J, Pilch A, Mury D, *et al*. Infliximab and adalimumab use during breastfeeding. *J Clin Gastroenterol* 2012;46:718–9.
8 Ben-Horin S, Yavzori M, Katz L, *et al*. Adalimumab level in breast milk of a nursing mother. *Clin Gastroenterol Hepatol* 2010;8:475–6.
9 Keeling S, Wolbink GJ. Measuring multiple etanercept levels in the breast milk of a nursing mother with rheumatoid arthritis. *J Rheumatol* 2010;37:1551.
10 Förger F, Zbinden A, Villiger PM. Certolizumab treatment during late pregnancy in patients with rheumatic diseases: Low drug levels in cord blood but possible risk for maternal infections. A case series of 13 patients. *Joint Bone Spine* 2016;83:341–3.
11 Mahadevan U, Wolf DC, Dubinsky M, *et al*. Placental transfer of anti-tumor necrosis factor agents in pregnant patients with inflammatory bowel disease. *Clin Gastroenterol Hepatol* 2013;11:286–92quiz e24.

12 Götestam Skorpen C, Hoeltzenbein M, Tincani A, *et al*. The EULAR points to consider for use of antirheumatic drugs before pregnancy, and during pregnancy and lactation. *Ann Rheum Dis* 2016;75:795–810.
13 Lawrence RM, Lawrence RA. Breastfeeding: more than just good nutrition. *Pediatr Rev* 2011;32:267–80.
14 Johnston M, Landers S, Noble L, *et al*. Breastfeeding and the use of human milk. *Pediatrics* 2012;129:e827–41.
15 Office of the Surgeon G, Centers for Disease C, Prevention and Office on Women's H. Publications and Reports of the Surgeon General. *The Surgeon General's Call to Action to Support Breastfeeding*. Rockville (MD): Office of the Surgeon General (US), 2011.
16 Hauck FR, Thompson JM, Tanabe KO, *et al*. Breastfeeding and reduced risk of sudden infant death syndrome: a meta-analysis. *Pediatrics* 2011;128:103–10.
17 Ip S, Chung M, Raman G, *et al*. Breastfeeding and maternal and infant health outcomes in developed countries. *Evid Rep Technol Assess* 2007:1–186.
18 Zelikin AN, Ehrhardt C, Healy AM. Materials and methods for delivery of biological drugs. *Nat Chem* 2016;8:997–1007.
19 Keystone E, Heijde D, Mason D, *et al*. Certolizumab pegol plus methotrexate is significantly more effective than placebo plus methotrexate in active rheumatoid arthritis: findings of a fifty-two-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. *Arthritis Rheum* 2008;58:3319–29.
20 Sandborn WJ, Feagan BG, Stoinov S, *et al*. Certolizumab pegol for the treatment of crohn's disease. *N Engl J Med* 2007;357:228–38.
21 Landewé R, Braun J, Deodhar A, *et al*. Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled phase 3 study. *Ann Rheum Dis* 2014;73:39–47.
22 Mease PJ, Fleischmann R, Deodhar AA, *et al*. Effect of certolizumab pegol on signs and symptoms in patients with psoriatic arthritis: 24-week results of a phase 3 double-blind randomised placebo-controlled study (RAPID-PsA). *Ann Rheum Dis* 2014;73:48–55.
23 Israel EJ, Taylor S, Wu Z, *et al*. Expression of the neonatal fc receptor, FcRn, on human intestinal epithelial cells. *Immunology* 1997;92:69–74.
24 Food and Drug Administration. *Guidance for industry: clinical lactation studies-study design, data analysis, and recommendations for labeling*, 2005.
25 Smeraglia J, Silva JP, Jones K. Improving the sensitivity and specificity of a bioanalytical assay for the measurement of certolizumab pegol. *Bioanalysis* 2017.
26 Dudal S, Baltrukonis D, Crisino R, *et al*. Assay formats: recommendation for best practices and harmonization from the global bioanalysis consortium harmonization team. *Aaps J* 2014;16:194–205.
27 Kruse N, Schulz-Schaeffer WJ, Schlossmacher MG, *et al*. Development of electrochemiluminescence-based singleplex and multiplex assays for the quantification of α-synuclein and other proteins in cerebrospinal fluid. *Methods* 2012;56:514–8.
28 Lacroix BD, Parker GL. S1029 dosing with Certolizumab Pegol (CZP) 200 mg every 2 weeks (Q2w) provides higher plasma trough concentrations than 400 mg every 4 weeks (Q4w). *Gastroenterology* 2010;138:S-163–4.
29 Wade JR, Parker G, Kosutic G, *et al*. Population pharmacokinetic analysis of certolizumab pegol in patients with Crohn's disease. *J Clin Pharmacol* 2015;55:866–74.
30 Alvares RD, Hasabnis A, Prosser RS, *et al*. Quantitative detection of PEGylated biomacromolecules in biological fluids by NMR. *Anal Chem* 2016;88:3730–8.
31 Hale T, Hartmann PE. *Textbook of human lactation*. Amarillo, Texas: Hale Publishing, 2007.
32 Bennett PN, Astrup-Jensen A, Bates CJ, *et al*. *Drugs and human lactation: a Comprehensive Guide to the content and consequences of drugs, micronutrients, radiopharmaceuticals and environmental and occupational chemicals in human milk*. 2nd ed: Elsevier Science & Technology, 1996.
33 Bykerk VP, Cush J, Winthrop K, *et al*. Update on the safety profile of certolizumab pegol in rheumatoid arthritis: an integrated analysis from clinical trials. *Ann Rheum Dis* 2015;74:96–103.
34 Cotton MF, Innes S, Jaspan H, *et al*. Management of upper respiratory tract infections in children. *South African Family Practice* 2008;50:6–12.
35 Worrall G. Common cold. *Can Fam Physician* 2011;57:1289–90.
36 Krol DM, Keels MA. Oral conditions. *Pediatr Rev* 2007;28:15–22.
37 Stengel JZ, Arnold HL. Is infliximab safe to use while breastfeeding? *World J Gastroenterol* 2008;14:3085–7.
38 Ostensen M, Eigenmann GO. Etanercept in breast milk. *J Rheumatol* 2004;31:1017–8.
39 Sandborn WJ, Lee SD, Randall C, *et al*. Long-term safety and efficacy of certolizumab pegol in the treatment of Crohn's disease: 7-year results from the PRECiSE 3 study. *Aliment Pharmacol Ther* 2014;40:903–16.
40 Rowe H, Baker T, Hale TW. Maternal medication, drug use, and breastfeeding. *Child Adolesc Psychiatr Clin N Am* 2015;24:1–20.
41 Elliott AB, Chakravarty EF. Management of rheumatic diseases during pregnancy. *Postgrad Med* 2010;122:213–21.
42 Neu J, Koldovsky O. Nutrient absorption in the preterm neonate. *Clin Perinatol* 1996;23:229–43.

EXTENDED REPORT

Mapping and predicting mortality from systemic sclerosis

Muriel Elhai,[1] Christophe Meune,[2] Marouane Boubaya,[3] Jérôme Avouac,[1] Eric Hachulla,[4] Alexandra Balbir-Gurman,[5] Gabriela Riemekasten,[6] Paolo Airò,[7] Beatriz Joven,[8] Serena Vettori,[9] Franco Cozzi,[10] Susanne Ullman,[11] László Czirják,[12] Mohammed Tikly,[13] Ulf Müller-Ladner,[14] Paola Caramaschi,[15] Oliver Distler,[16] Florenzo Iannone,[17] Lidia P Ananieva,[18] Roger Hesselstrand,[19] Radim Becvar,[20] Armando Gabrielli,[21] Nemanja Damjanov,[22] Maria J Salvador,[23] Valeria Riccieri,[24] Carina Mihai,[25] Gabriella Szücs,[26] Ulrich A Walker,[27] Nicolas Hunzelmann,[28] Duska Martinovic,[29] Vanessa Smith,[30] Carolina de Souza Müller,[31] Carlo Maurizio Montecucco,[32] Daniela Opris,[33] Francesca Ingegnoli,[34] Panayiotis G Vlachoyiannopoulos,[35] Bojana Stamenkovic,[36] Edoardo Rosato,[37] Stefan Heitmann,[38] Jörg H W Distler,[39] Thierry Zenone,[40] Matthias Seidel,[41] Alessandra Vacca,[42] Ellen De Langhe,[43] Srdan Novak,[44] Maurizio Cutolo,[45] Luc Mouthon,[46] Jörg Henes,[47] Carlo Chizzolini,[48] Carlos Alberto von Mühlen,[49] Kamal Solanki,[50] Simona Rednic,[51] Lisa Stamp,[52] Branimir Anic,[53] Vera Ortiz Santamaria,[54] Maria De Santis,[55] Sule Yavuz,[56] Walter Alberto Sifuentes-Giraldo,[57] Emmanuel Chatelus,[58] Jiri Stork,[59] Jacob van Laar,[60] Esthela Loyo,[61] Paloma García de la Peña Lefebvre,[62] Kilian Eyerich,[63] Vanesa Cosentino,[64] Juan Jose Alegre-Sancho,[65] Otylia Kowal-Bielecka,[66] Grégoire Rey,[67] Marco Matucci-Cerinic,[68] Yannick Allanore,[1] on behalf of EUSTAR group

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-2017-211448).

For numbered affiliations see end of article.

Correspondence to
Professor Yannick Allanore, Paris Descartes University Cochin Hospital, Rheumatology A department, 27 rue du Faubourg Saint Jacques 75014 Paris, France; yannick.allanore@cch.aphp.fr

Received 9 March 2017
Revised 21 May 2017
Accepted 18 July 2017
Published Online First 23 August 2017

To cite: Elhai M, Meune C, Boubaya M, *et al*. *Ann Rheum Dis* 2017;**76**:1897–1905.

ABSTRACT

Objectives To determine the causes of death and risk factors in systemic sclerosis (SSc).

Methods Between 2000 and 2011, we examined the death certificates of all French patients with SSc to determine causes of death. Then we examined causes of death and developed a score associated with all-cause mortality from the international European Scleroderma Trials and Research (EUSTAR) database. Candidate prognostic factors were tested by Cox proportional hazards regression model by single variable analysis, followed by a multiple variable model stratified by centres. The bootstrapping technique was used for internal validation.

Results We identified 2719 French certificates of deaths related to SSc, mainly from cardiac (31%) and respiratory (18%) causes, and an increase in SSc-specific mortality over time. Over a median follow-up of 2.3 years, 1072 (9.6%) of 11 193 patients from the EUSTAR sample died, from cardiac disease in 27% and respiratory causes in 17%. By multiple variable analysis, a risk score was developed, which accurately predicted the 3-year mortality, with an area under the curve of 0.82. The 3-year survival of patients in the upper quartile was 53%, in contrast with 98% in the first quartile.

Conclusion Combining two complementary and detailed databases enabled the collection of an unprecedented 3700 deaths, revealing the major contribution of the cardiopulmonary system to SSc mortality. We also developed a robust score to risk-stratify these patients and estimate their 3-year survival. With the emergence of new therapies, these important observations should help caregivers plan and refine the monitoring and management to prolong these patients' survival.

INTRODUCTION

Systemic sclerosis (SSc) is a devastating disease that has a profound impact on life expectancy, reflected by a standardised mortality ratio of 3.5.[1] Its discordant causes and predictors of death have been studied in mostly small samples from single institutions, limiting their application to new studies of epidemiology.[1–10] Because the presentation and prognosis of SSc are highly heterogeneous, the identification of patients at high risk of death, who may benefit from close monitoring and early treatment, is crucial.

Among various methods available to determine the causes of death, the analysis of death certificates is considered robust,[11] although it has been scarcely used in investigations of SSc, with no report after year 2000.[12] The ongoing European Scleroderma Trials and Research (EUSTAR) is an international, multicentre, prospective registry managed by physicians (list of authors and online supplementary

appendix 1) and organised centrally by its committee.[13] This database offers a unique opportunity to study the natural history of the disease and predict outcomes through the prospective, standardised collection of multiple characteristics of patients with SSc. Since the first report based on 284 deaths among 5860 patients in 2010,[14] the database has grown to >11 000, and the numbers of follow-up visits and deaths have increased accordingly.

Our aim was to identify the specific causes of death and their respective incidence by reviewing all death certificates of patients presenting with SSc, collected in France between 2000 and 2011, using a multiple-cause-of-death analysis.[15 16] We then examined the causes of death and associated factors to develop a risk score associated with overall mortality in the international EUSTAR sample.

METHODS

Death certificates

All death certificates issued in France comply with the international standards of the WHO and are exhaustively collected by the '*Centre d'épidémiologie sur les causes médicales de décès*' (Epidemiological Centre for the Medical Causes of Death — CépiDc) from the '*Institut national de la santé et de la recherche médicale*' (National Health and Medical Research Institute — INSERM).[17] In January 2015, we examined the certificates of all adults presenting with SSc (international classification of diseases (ICD)-10 code M34) who died between 1 January 2000 and 31 December 2011.

Table 1 Absolute number of deaths related to systemic sclerosis in France between 2000 and 2011

All systemic sclerosis-related deaths	2719
Systemic sclerosis listed as underlying cause of death	1608
Females	1276
Males	332
Female/male ratio	3.8
Age, year	
<50	119
50–59	171
60–69	330
70–79	544
>80	444
Systemic sclerosis listed as associated cause of death	1111
Females	881
Males	230
Female/male ratio	3.8
Age, year	
<50	65
50–59	99
60–69	211
70–79	388
>80	348
Age-standardised mortality rate	
All patients presenting with systemic sclerosis	0.80
Females	1.03
Males	0.41
Female/male ratio	2.49

Unless indicated otherwise, values are raw counts.

Statistical analysis

A multiple-cause-of-death analysis was performed allowing the retrieval of the death certificates, which listed SSc as the 'underlying' cause of death (UCD) and those which considered SSc as the 'associated' cause of death (ACD).[16 18 19]

Mortality rates were calculated by age group for the entire period from 2000 to 2011. Age-standardised mortality rates per 10^5 patients were calculated by a direct method, per year and for the study period, using the standard 2000–2011 population data of the European Union and the European Free Trade Association.

To measure the strength of association between SSc and the various causes of death, we calculated the observed number of deaths in relation to the expected number of deaths (O/E ratio), based on the proportional mortality rate for the same cause of death within the French general population between 2000 and 2011. An O/E >1 means an excess mortality associated with SSc.

The EUSTAR sample

We interrogated the EUSTAR database at the end of May 2014, providing information on 11 193 patients >18 years age, from 124 participating centres, fulfilling the 2013 criteria formulated for SSc by the American College of Rheumatology/European League Against Rheumatism.[20] The structure of the database, the minimum essential data set and the inclusion criteria have been described in detail previously.[13] Each participating centre obtained approval of the local ethics committee and all registered patients granted their informed consent. Among the 11 193 patients who underwent ≥1 visit, 7819 had ≥1 follow-up and 1072 died. Besides the disease characteristics and treatment, we recorded the date of death and whether the death was attributable to SSc or to another cause. Furthermore, we probed the participating centres with a view to identify a single pulmonary,

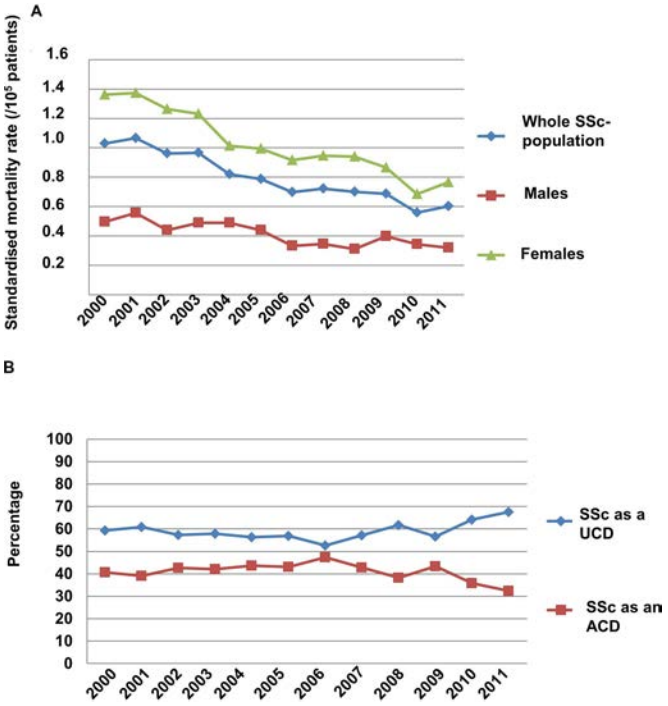


Figure 1 Deaths and systemic sclerosis in France between 2000 and 2011. (A) Age-standardised mortality per 10^5 men, women or both. (B) Percentage of deaths among patients presenting with systemic sclerosis (SSc) as the underlying cause of death (UCD) versus an associated cause of death (ACD).

Table 2 (A) Sex-adjusted and age-adjusted O/E ratios for the cause of death in SSc and (B) age-adjusted O/E ratios for the cause of death in SSc

(A)

	Male			Female			Total		
	Patients with SSc n=562	General population n=3 327 105	O/E ratio (95% CI)	Patients with SSc n=2157	General population n=3 168 152	O/E ratio (95% CI)	Patients with SSc n=2719	General population n=6 495 257	O/E ratio (95% CI)
Cardiovascular disease	200 (35.6%)	855 720 (25.7%)	1.38 (1.24 to 1.55)	870 (40.3%)	980 413 (30.9%)	1.30 (1.24 to 1.37)	1070 (39.3%)	1 836 133 (28.3%)	1.39 (1.33 to 1.46)
Respiratory disease	135 (24.0%)	211 796 (6.4%)	3.77 (3.28 to 4.37)	371 (17.2%)	192 326 (6.1%)	2.83 (2.58 to 3.11)	506 (18.6%)	404 722 (6.2%)	2.99 (2.76 to 3.23)
Infectious disease	59 (10.5%)	65 069 (1.9%)	5.37 (4.22 to 6.83)	244 (11.3%)	63 957 (2.0%)	5.60 (4.98 to 6.31)	303 (11.1%)	129 026 (2.0%)	5.61 (5.04 to 6.24)
Malignant neoplasm	75 (13.3%)	1 077 631 (32.3%)	0.41 (0.33 to 0.51)	179 (11.9%)	735 610 (23.2%)	0.36 (0.31 to 0.41)	254 (9.3%)	1 813 241 (27.9%)	0.33 (0.30 to 0.38)
Gastrointestinal disease	28 (5.0%)	154 423 (4.6%)	1.07 (0.75 to 1.54)	151 (7.0%)	132 897 (4.2%)	1.68 (1.44 to 1.96)	179 (6.6%)	287 320 (4.4%)	1.49 (1.29 to 1.71)

(B)

	<60 years			60–79 years			≥80 years		
	Patients with SSc n=454	General population n=1 020 794	O/E ratio (95% CI)	Patients with SSc n=1473	General population n=2 162 597	O/E ratio (95% CI)	Patients with SSc n=792	General population n=3 311 866	O/E ratio (95% CI)
Cardiovascular disease	176 (38.8%)	125 895 (12.3%)	3.14 (2.80 to 3.53)	576 (39.1%)	524 385 (24.2%)	1.61 (1.51 to 1.72)	318 (40.1%)	1 185 853 (35.8%)	1.12 (1.03 to 1.22)
Respiratory disease	94 (20.7%)	22 248 (2.2%)	9.50 (7.93 to 11.38)	292 (19.8%)	113 894 (5.3%)	3.76 (3.40 to 4.17)	120 (15.1%)	268 580 (8.1%)	1.87 (1.58 to 2.20)
Infectious disease	56 (12.3%)	23 135 (2.3%)	5.44 (4.26 to 6.96)	162 (11.0%)	40 009 (1.8%)	5.94 (5.14 to 6.88)	85 (10.7%)	65 882 (2.0%)	5.39 (4.41 to 6.60)
Malignant neoplasm	60 (13.2%)	362 804 (35.5%)	0.37 (0.29 to 0.47)	135 (9.2%)	881 509 (40.8%)	0.22 (0.19 to 0.26)	58 (7.3%)	568 928 (17.2%)	0.43 (0.33 to 0.55)
Gastrointestinal disease	30 (6.6%)	61 906 (6.1%)	1.09 (0.77 to 1.54)	95 (6.4%)	102 975 (4.8%)	1.35 (1.11 to 1.64)	27 (3.4%)	122 439 (3.7%)	0.92 (0.64 to 1.34)

Unless specified otherwise, the values are numbers (%) of observations.
O/E, observed/expected; SSC, systemic sclerosis.

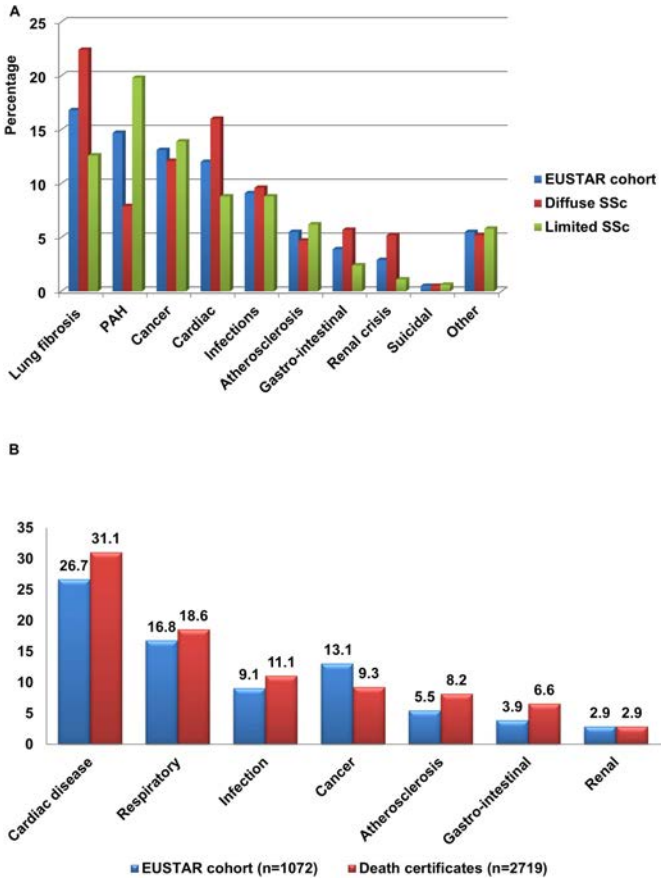


Figure 2 (A) Causes of death in the entire EUSTAR sample and in the limited and diffuse cutaneous forms. (B) Comparison of causes of death in the EUSTAR and in the death certificates samples. The results are presented as % of deaths. EUSTAR, European Scleroderma Trials and Research; PAH, pulmonary arterial hypertension; SSc, systemic sclerosis.

cardiac, renal, infectious, neoplastic, gastrointestinal, suicidal or other primary cause of death, according to a standard set of definitions, and to record any clinically significant comorbidity in a brief additional form submitted to all centres where ≥1 patient death was entered in the database (online supplementary appendix 2).

Statistical analysis

Categorical results are presented as counts and percentages, and continuous variables as mean±SD.

Survival and prognostic score

The median (95% CI) follow-up was estimated by the reverse Kaplan-Meier method, and the overall survival by the Kaplan-Meier method. Potential prognostic factors were analysed first by the Cox proportional hazards regression model in single variable analysis. The proportional hazards assumption was verified by Schoenfeld residuals.[21] Continuous variables were dichotomised according to the clinical cut-off.

To ascertain a possible linearity among the variables, the variance inflation factor was calculated, and the variables were considered colinear when >2.[22] All factors emerging with p values <0.10 by single variable analysis were included in a multiple variable model and stratified by centre. Due to the multicollinearity and missing data for the former, the cutaneous form of the disease and muscle weakness were selected instead of the Rodnan score and muscle atrophy, respectively.

A backward, stepwise variable selection algorithm was applied using a stopping rule based on a cut-off p value of 0.05. To account for missing observations, the data were analysed, using multiple imputations by chained equations, with 50 imputations obtained after 20 iterations.[23 24] The variables considered in the imputation models were all the characteristics studied as prognostic factors, death status and Nelson-Aalen estimator of the cumulative hazard. In these variables, missing values ranged from 0% to 56.5%, with a median value of 2.0%. The results were aggregated by pooling the estimates obtained on each imputed data set according to Rubin's rules. To develop the SCleroderma mOrtality p Eustar (SCOpE) prognostic score to use in clinical practice, we assigned points by rounding the beta values multiplied by 5 for the significant predictors, in order to obtain a minimal factor of 1.

The discriminative ability of the models was evaluated by the C-index after bootstrap correction for overoptimism, and by receiver operating characteristics (ROC) curve and area under the curve (AUC) for 3-year mortality. The models calibration was assessed by the calibration slope and the bootstrap, bias-corrected calibration slope at 3 years. The

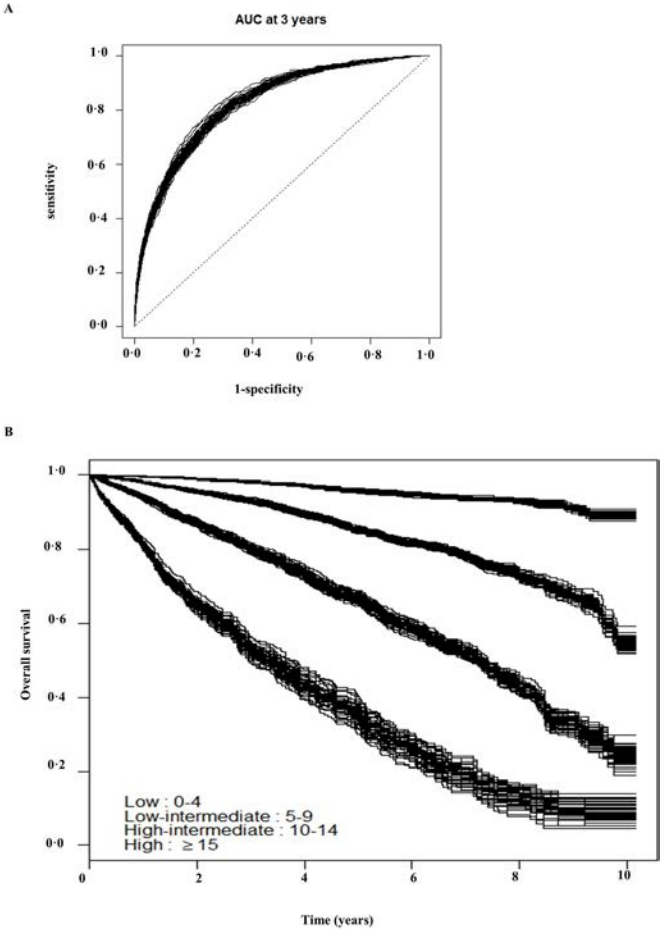


Figure 3 (A) ROC curves at 3 years for the SCOpE score. The lines indicate ROC curves of 50 models from 50 imputed data sets. (B) Overall survival according to simplified score categories. Curves are plotted for each of the 50 imputed data set. Three-year survival according to SCOpE score: 0.98 (0.97–0.99) (score: 0–4); 0.93 (0.92–0.94) (score: 5–9); 0.80 (0.78–0.83) (score: 10–14); 0.53 (0.48–0.58) (score ≥15). AUC, area under the curve; ROC, receiver operating characteristics; SCOpE, SCleroderma mOrtality p Eustar.

overoptimism induced by the models was corrected by multiplying the regression coefficients by the calibration slope.[22] The 50 imputed data sets were internally validated by bootstrapping with random generation of 200 samples from the original data. This score was compared with the previous Bryan score using an ROC analysis.[7]

All tests were two-sided at a 0.05 significance level. The analyses were carried out using the R V.3.1.2 statistical software. Further details are in the online supplementary methods.

RESULTS

Death certificates

Causes of death

Between 2000 and 2011, 6 474 953 adults died in France. SSc was listed in 2719 death certificates, including 1608 as UCD and 1111 as ACD, representing 0.04% of all death certificates issued during the study period (table 1). The mean age at the time of death was 71.4±12.8 years (online supplementary figure 1). The female/male (2157 female and 562 male) sex ratio was 3.8. The causes of death were cardiac in 31%, respiratory in 18%, infectious in 11% and cancers in 9% of cases (online supplementary table 1). Further information is in the online supplementary information.

Mortality trends between 2000 and 2011

The overall, age-standardised mortality rate among patients with SSc was 0.80 per 10^5 individuals, with a female/male ratio of 2.49 (table 1). This rate decreased gradually from 1.03 per 10^5 men and women in year 2000, to 0.60 per 10^5 in year 2011 (figure 1A). The female-to-male ratio remained stable throughout the period. The ratio of deaths in which SSc was the UCD increased between 2000 and 2011, whereas the proportion of deaths in which SSc was the ACD decreased (figure 1B).

Comparison of causes of death with the general population

The O/E ratios for cardiovascular, respiratory and infectious diseases were 1.36, 2.99 and 5.61, respectively, whereas the O/E for malignancy was 0.33. The excess mortality associated with respiratory

Table 3 Predictors of low survival in the multiple variable model

	Mode				
	Full		Final		Simplified score
	HR (95% CI)	p	HR (95% CI) p	p	
Age, year					
50–65	1.93 (1.6 to 2.32)	<0.001	1.86 (1.56 to 2.21)	<0.001	3
>65	3.91 (3.2 to 4.78)	<0.001	3.63 (3.02 to 4.38)	<0.001	6
Male sex	1.37 (1.15 to 1.64)	<0.001	1.34 (1.13 to 1.58)	0.001	1
Diffuse cutaneous disease	0.79 (0.67 to 0.93)	0.006	1.25 (1.08 to 1.46)	0.004	1
>5 years disease duration	0.91 (0.79 to 1.06)	0.23	–	–	–
Progressive digital vasculopathy*	0.91 (0.66 to 1.27)	0.58	–	–	–
Oesophageal or gastric disease manifestations	1.04 (0.88 to 1.23)	0.65	–	–	–
Intestinal involvement	1.13 (0.97 to 1.32)	0.13	–	–	–
Systemic hypertension	0.94 (0.79 to 1.11)	0.46	–	–	–
Scleroderma renal crisis	1.56 (1.05 to 2.32)	0.029	1.48 (1.02 to 2.15)	0.039	2
Palpitations	1.14 (0.97 to 1.35)	0.12	–	–	–
Prominent dyspnoea	1.81 (1.41 to 2.31)	<0.001	1.79 (1.43 to 2.24)	<0.001	3
Digital ulcers	1.27 (1.1 to 1.47)	0.001	1.24 (1.08 to 1.42)	0.002	1
Joint synovitis	1.00 (0.83 to 1.21)	0.98	–	–	–
Contracture	1.3 (1.1 to 1.52)	0.002	1.28 (1.1 to 1.49)	0.001	1
Tendon friction rub	0.96 (0.77 to 1.21)	0.75	–	–	–
Muscle weakness	1.3 (1.1 to 1.54)	0.002	1.34 (1.14 to 1.56)	<0.001	1
Elevated C reactive protein	2.47 (1.93 to 3.15)	<0.001	2.34 (1.88 to 2.93)	<0.001	4
Elevated creatine kinase	1.09 (0.86 to 1.38)	0.49	–	–	–
Proteinuria	2.04 (1.59 to 2.61)	<0.001	1.95 (1.53 to –2.47)	<0.001	3
Left ventricular ejection fraction <50%	1.46 (1.07 to –2.01)	0.019	1.41 (1.04 to 1.91)	0.027	2
Pulmonary arterial hypertension*†	1.13 (0.65 to 1.95)	0.67	–	–	–
Interstitial lung disease	1.28 (1.09 to 1.5)	0.003	1.26 (1.08 to 1.46)	0.003	1
Carbon monoxide diffusion capacity <60% predicted	2.07 (1.75 to 2.44)	<0.001	2.02 (1.72 to 2.38)	<0.001	4
Forced vital capacity <70% predicted	1.41 (1.13 to 1.76)	0.003	1.4 (1.13 to 1.73)	0.002	2
Disease activity score =3	0.85 (0.63 to 1.14)	0.28	–	–	–
Antinuclear antibodies	1.04 (0.76 to 1.45)	0.79	–	–	–
Anti-Scl70 antibodies	0.98 (0.83 to 1.16)	0.8	–	–	–

*In the last month, dyspnoea was classified as prominent in presence of New York Heart Association functional class III or IV.
†Diagnosed at time of right heart catheterisation; interstitial lung disease was considered present if visible on chest radiograph or on high-resolution CT scan; disease was active if the disease activity score was ≥3; the full model contains all variables included in the multiple variable model. The final model is model after variable selection. The HRs are pooled over the 50 imputed data sets and divided by the calibration slope of 0.94. Simplified score points were attributed to the variables of the final model by rounding the regression coefficients multiplied by 5.

diseases (O/E=3.77) was particularly prominent in men, while that associated with cardiovascular (O/E=3.14) and respiratory (O/E=9.50) diseases strongly involved patients <60 years (table 2).

EUSTAR sample

Causes of death

A total of 11 193 patients with SSc were identified in the EUSTAR sample (online supplementary table 3). Of these, 86% were women, 31.0% presented with the diffuse cutaneous subtype and the mean disease duration was 8.1 years. Of these patients, 1072 (9.6%) died. The mean age at time of death was 63.6±13.4 years and the mean disease duration was 12.3±12.4 years (online supplementary figure 2). Death was considered SSc-related in 617 cases (57.6%) and unrelated to SSc in 270 cases (25.2%).

Additional forms were completed for 940/1072 (87.7%) deaths by 64 participating centres (figure 2 and online supplementary table 3). The main causes of death were interstitial lung disease (ILD) (16.8%), pulmonary arterial hypertension (PAH) (14.7%), cancer (13.1%), primary heart disease (12.0%) and infection (9.1%). Further details are in the online supplementary information.

Predictors of death and prognostic score

Among 11 193 patients entered in the database, 7819 had ≥1 additional follow-up after the first visit (median follow-up: 2.3 (1.3–5.3) years). The disease characteristics of the patients with versus without ≥1 additional follow-ups were significantly dissimilar by single, though not by multiple variable analysis (online supplementary table 5). The 3-year survival rate (online supplementary figure 3) was 89.3% (88.5%–90.2%). The 39 variables associated with the 3-year mortality by single variable analysis are listed in online supplementary table 6. Online supplementary table 7 shows the description of the full model variables (1) according to the original data set (without imputation) and (2) averaged over all complete data sets (including the imputed data). No significant difference was observed between the two models. By Cox multiple variable regression analysis, age, male sex, the cutaneous subset of the disease, elevated C reactive protein, class II–IV dyspnoea, ILD, low carbon dioxide diffusing capacity, forced vital capacity, proteinuria, scleroderma renal crisis, depressed left ventricular ejection fraction, digital ulcers and joint involvement were independent predictors of 3-year mortality (table 3), allowing the development of the SCOpE score, ranging between 0 and 32. With an average corrected C-index of 0.80, this score was discriminate. At 3 years, the average AUC was 0.82 (95% CI 0.80 to 0.84; figure 3A). The AUC for 3-year mortality was 0.79 (95% CI 0.75 to 0.81) for diffuse and 0.82 (95% CI 0.80 to 0.85) for limited SSc (online supplementary figure 3c). This score was discriminative for both incident (<1 year) and prevalent SSc (online supplementary figure 3D). The discrimination power of the SCOpE score for 3-year mortality was higher (AUC of 0.82 (95% CI 0.80 to 0.84)) than that of the Bryan score (AUC 0.72 (95% CI 0.70 to 0.74); p<0.001; online supplementary figure 3E).[7 8] When divided into quartile, 599 patients with scores ≥15 had a 0.53 (95% CI 0.48 to 0.58) 3-year survival rate, compared with 2777 patients with scores <5, whose 3-year survival rate was 0.98 (95% CI 0.97 to 0.99) (p<0.001; figure 3B).

DISCUSSION

The strengths of our report include our two-step study with first the collection of all death certificates in France during a 10-year period, corresponding to the analysis of 2719 death certificates from patients with SSc, followed by the interrogation of the very large EUSTAR database that included 11 193 patients and 1072 deaths at the time of the analysis. This large collection of patients represents the most robust report of any mortality study and prediction score. Our analysis of two distinct sources of information and the consistency of our results are evidence that our methodology mitigated the effects of common biases observed in previous studies.

We confirmed that primary heart disease is the main offender in SSc explaining 30% of SSc deaths,[1 4 6 14 25 26] while atherosclerosis was responsible for only 5%–8% of deaths.[26] This highlights the importance of thorough cardiac investigations to identify patients presenting with SSc at a preclinical stage of PAH and cardiac involvement. Except for systemic hypertension, neither the EUSTAR sample nor the death certificates included a list of cardiovascular risk factors, preventing a correction of the causes of deaths for rates of risk factors. However, in a previous EUSTAR study, the typical cardiovascular risk factors were not identified as important contributors to heart involvement.[27]

We confirmed that lung involvement is a major complication of SSc, particularly in young patients and in men who, compared with the general population, suffered respectively tenfold and fourfold higher rates of deaths from respiratory diseases. Accordingly, respiratory failure was recently shown to contribute prominently to intensive care unit admissions for management of SSc.[28] Besides the high mortality associated with respiratory failure, our study revealed a high mortality from lung infections and a fivefold higher rate of infectious deaths among patients with SSc compared with the general population. These observations highlight the importance of the infectious risk associated with this disease and of the need to use specific therapeutic measures that are underused, such as vaccinations.[29]

We also observed a high proportion of death from cancer, of the lung in particular, although compared with the general population, the risk of death from cancer was not increased, in contrast to other autoimmune diseases.[30] Alternatively, premature death due to terminal SSc may have obscured the age-related increase in deaths from cancer. Finally, the death certificates might have failed to mention the diagnosis of SSc when patients died from cancer.

We observed a gradual decrease in standardised mortality rate over time due to a decrease in mortality unrelated to SSc, while the rate of deaths due to SSc increased. One possible explanation for this finding is that increased survival among the general population may largely account for the increased survival observed in SSc in this study.[1] These observations should encourage the community to urgently revise and improve the care of SSc, by focusing on a more accurate identification of poor-prognosis patients, who might benefit from aggressive therapy, and from the development of a critically needed reliable prognostic score.

For this purpose, we developed a weighted risk equation for survival at 3 years from a sample of over 11 000 patients, based on a rigorous data collection by study centres highly skilled in SSc management. There was only a median of 2.0% of missing prognostic variables and we used imputations to minimise the possible role of missing values, and stratified the data analysis by study centre. The respective weight of the selected variables was similar before and after imputation (online supplementary table 7), confirming the robustness of our sample and of our data collection. The AUC of the SCOpE to predict the 3-year mortality was 0.82, and the reliability of our score was confirmed by bootstrapping analysis. This SCOpE ranged from 0 to 32, and is simple to calculate (online supplementary appendix 3). When compared with

the Bryan score, our SCOpE score was more discriminate ($p<0.001$). This score confirmed its robustness in incident and prevalent SSc, as well as in limited and diffuse cutaneous subtypes, suggesting that it is applicable to all patients presenting with this disease. Using that score, we were able to stratify patients among four sharply distinct groups of severity. This risk stratification might help adapt the monitoring to the specific risk represented by a patient, contribute to decision for expert centre referral and advance the diagnosis of internal organ involvement in patients whose score is ≥15. Furthermore, the SCOpE score might help select the candidates for high-level therapeutic interventions, such as stem cell transplantation, and for inclusion in clinical trials and preventive strategies. These broad applications should be validated in dedicated studies. However, our study should be interpreted within its limitations: (1) the precise cause of death may be difficult to ascertain, for example in patients who died away from diagnostic facilities. This may explain disparities between death certificates and adjudicated expert judgement.[31] For example, in death certificates, pulmonary embolism was believed to be responsible for 1/3 of cardiac-related deaths. We can hypothesise that most of these deaths might be secondary to right heart involvement or PAH, which were under-recognised by non-experts in SSc. The absence of detailed clinical records and information regarding concomitant illnesses may also bias the death certificates, although the inclusion of a large number of certificates in the analysis should mitigate such biases. Furthermore, our observation of similar causes of death in the certificates analysis and in the EUSTAR sample supports our methodology. (2) The mean disease duration in EUSTAR cohort was over 8 years, which might cause missing of early deaths. However, thanks to the very large population included, we assume this cohort is a representation of our current practice. In addition, early SSc (<3 years) was not associated with mortality. (3) We were not able to externally validate the final model, but we have used the bootstrapping method as a validation tool. Bootstrapping is a robust method that is thought to be used when no external cohort of patients is available.[32] (4) Three thousand patients did not have at least one follow-up visit. The disease characteristics were not significantly different in multivariate analysis between patients with and without follow-up, which suggests that it may not have influenced our results. (5) We decided to not include the treatments in our prediction model because (1) in the absence of strict recommendations, many of the disparities observed are based on clinical considerations instead of various forms of the disease, and (2) we wished to develop a score applicable to new patients as well as patients already treated. (6) Finally, since both our study samples included Caucasians, our score cannot be extrapolated to other ethnic groups.

To conclude, our study should impress the community by the lack of progress it reveals in the survival of patients with SSc. An early and systematic management of the large proportion of cardiac complications associated with this disease is in order, in hope of extending survival in SSc. Because of the large difference in mortality compared with the general population, lung involvement as well as infections should be prominently visible on the research agenda. We also developed a robust mortality score to estimate the 3-year survival and risk-stratify patients. With the emergence of new therapies in SSc, these results should help caregivers adapt the monitoring and therapeutic strategies to the specific risk of each patient, with a view to prolong the survival in SSc.

Author affiliations
[1]Rheumatology A department, Paris Descartes University, INSERM U1016, Sorbonne Paris Cité, Cochin Hospital, Paris, France
[2]Department of Cardiology, Paris XIII University, INSERM UMR S-942, Bobigny Hospital, Paris, France
[3]Unit of Clinical Research, Paris Seine Saint Denis University, Bobigny, France
[4]Department of Internal Medicine, Hôpital Claude Huriez, University Lille Nord-de-France, Lille Cedex, Lille, France
[5]B Shine Rheumatology Unit, Rambam Health Care Campus, Rappaport Faculty of Medicine, Technion—Institute of Technology, Haifa, Israel
[6]Department of Rheumatology, University of Lübeck, Lübeck, Germany
[7]UO Reumatologia ed Immunologia Clinica Spedali Civili Brescia, Brescia, Italy
[8]Servicio de Reumatologia, Hospital Universitario 12 de Octubre, Madrid, Spain
[9]Department of Clinical and Experimental Medicine, 'F-Magrassi' II, Naples, Italy
[10]Rheumatology Unit, Department of Medicine, University of Padova, Padova, Italy
[11]Department of Dermatology, University Hospital of Copenhagen, Hospital Bispebjerg, Copenhagen, Denmark
[12]Department of Immunology and Rheumatology, University of Pécs, Pécs, Hungary
[13]Chris Hani Baragwanath Academic Hospital University of the Witwatersrand, Johannesburg, South Africa
[14]Department of Rheumatology and Clinical Immunology, Justus-Liebig University Giessen, Kerckhoff Clinic, Bad Nauheim, Germany
[15]Rheumatology Unit, Department of Medicine, University of Verona, Verona, Italy
[16]Department of Rheumatology, University Hospital Zurich, Zurich, Switzerland
[17]Interdisciplinary Department of Medicine-Rheumatology Unit, Policlinico, University of Bari, Bari, Italy
[18]VA Nasonova Institute of Rheumatology, Moscow, Russian Federation
[19]Department of Clinical Sciences Lund, Section of Rheumatology, Lund University, Skåne University Hospital, Lund, Sweden
[20]Institute of Rheumatology, 1st Medical School, Charles University, Praha, Czech Republic
[21]Clinica Medica, Dipartimento di Scienze Cliniche e Molecolari, Università Politecnica delle Marche, Ancona, Italy
[22]Institute of Rheumatology, University of Belgrade Medical School, Belgrade, Serbia
[23]Rheumatology Department, Hospitais da Universidade, Coimbra, Portugal
[24]Department of Internal Medicine and Medical Specialities, 'Sapienza', University of Rome, Italy, Rome, Italy
[25]Department of Internal Medicine and Rheumatology Clinic, Ion Cantacuzino Clinical Hospital, Carol Davila University of Medicine and Pharmacy, Bucharest, Romania
[26]Department of Internal Medicine, Division of Rheumatology, University of Debrecen, Debrecen, Hungary
[27]Department of Rheumatology, Basel University, Unispital Basel, Basel, Switzerland
[28]Department of Dermatology, University Hospital Cologne, Cologne, Germany
[29]Department of Internal Medicine, Clinical Hospital of Split, Split, Croatia
[30]Department of Rheumatology, University of Ghent, Ghent, Belgium
[31]Rheumatology Division, Clinics Hospital, Federal University of Paraná, Curitiba, Brazil
[32]Unita Operativa e Cattedra di Reumatologia, IRCCS Policlinico S Matteo, Pavia, Italy
[33]Department of Rheumatology, St Maria Hospital, Carol Davila, University of Medicine and Pharmacy, Bucharest, Romania
[34]Dipartimento e Cattedra di Reumatologia, Università degli Studi di Milano, Istituto Ortopedico 'Gaetano Pini', Milano, Italy
[35]Department of Pathophysiology, Medical School, National University and Kapodistrian of Athens, Athens, Greece
[36]Institute for Prevention, Treatment and Rehabilitation of Rheumatic and Cardiovascular Diseases, Niska Banja, Serbia and Montenegro
[37]Dipartimento di Medicina Clinica, Centro per la Sclerosi Sistemica, Università La Sapienza, Policlinico Umberto I, Roma, Italy
[38]Department of Rheumatology, Marienhospital Stuttgart, Stuttgart, Germany
[39]Department of Internal Medicine 3, University Hospital Erlangen, Erlangen, Germany
[40]Department of Medicine, Unit of Internal Medicine, Valence cedex, France
[41]Medizinische Klinik III, University Hospital of Bonn, Bonn, Germany
[42]Rheumatology Unit, University Hospital of Cagliari, Monserrato, Italy
[43]Division of Rheumatology and Department of Development and Regeneration, University Hospital Leuven and Laboratory Tissue Homeostasis and Disease, Leuven, Belgium
[44]Department of Rheumatology and Clinical Immunology, Internal Medicine, KBC Rijeka, Rijeka, Croatia
[45]Research Laboratory and Division of Rheumatology Department of Internal Medicine, University of Genova, Genova, Italy
[46]Department of Internal Medicine, Hôpital Cochin, Paris, France
[47]Medizinische Universitätsklinik, Abt II (Onkologie, Hämatologie, Rheumatologie, Immunologie, Pulmonologie), Tübingen, Germany
[48]Department of Immunology and Allergy, University Hospital, Geneva, Switzerland
[49]Rheuma Clinic Av Carlos Gomes Porto Alegre, Porto Alegre, Brazil

[50]Rheumatology Unit, Waikato University Hospital, Hamilton City, Hamilton, New Zealand
[51]Department of Rheumatology, University of Medicine and Pharmacy 'Iuliu Hatieganu' Cluj, Cluj-Napoca, Romania
[52]Department of Medicine, University of Otago, Christchurch, New Zealand
[53]Division of Clinical Immunology and Rheumatology, Department of Medicine, University Hospital Centre Zagreb, Zagreb, Croatia
[54]Rheumatology Granollers General Hospital, Barcelona, Spain
[55]Division of Rheumatology and Clinical Immunology, Humanitas Clinical and Research Center, Rozzano, Italy
[56]Department of Rheumatology, University of Marmara, Istanbul, Turkey
[57]Servicio de Reumatología, Hospital Ramon Y Cajal, Madrid, Spain
[58]Department of Rheumatology, University Hospital of Strasbourg-, Hôpital de Hautepierre, Service de Rhumatologie, Strasbourg, France
[59]Department of Dermatology, Charles University and General University Hospital in Prague, Prague, Czech Republic
[60]Department of Rheumatology and Clinical Immunology, James Cook University Hospital, Middlesbrough, UK
[61]Reumatologia e Inmunologia Clinica, Hospital Regional Universitario Jose Ma Cabral y Baez, Clinica Corominas, Santiago, Dominican Republic
[62]Rheumatology Department, Hospital Universitario Madrid Norte Sanchinarro, Madrid, Spain
[63]Department of Dermatology and Allergy of the TU Munich, Munich, Germany
[64]Department of Rheumatology and Collagenopathies, Osteoarticular Diseases and Osteoporosis Centre, Pharmacology and Clinical Pharmacological Research Centre, School of Medicine-University of Buenos Aires, Ramos Mejía Hospital, Buenos Aires, Argentina
[65]Department of Rheumatology, Hospital Universitario Dr Peset, Valencia, Spain
[66]Department of Rheumatology and Internal Medicine, Medical University of Bialystok, Bialystok, Poland
[67]INSERM, CépiDc, Le Kremlin- Bicêtre, Le Kremlin-Bicêtre, France
[68]Department of Experimental and Clinical Medicine, Section of Internal Medicine and Division of Rheumatology, Azienda Ospedaliero-Universitaria Careggi (AOUC), University of Florence, Florence, Italy

Contributors ME, CM and YA formulated the study hypotheses and contributed to its design and analysis of the data, literature search, composition of the tables and figures, and redaction of the first draft and subsequent iterations of the manuscript. ME, MB and CM performed the statistical analyses. JA, EH, ABG, GR, PA, BJ, SV, FC, SU, LC, MT, UML, PC, OD, FI, LPA, RH, RB, AG, ND, MJS, VR, CM, GS, UAW, NH, DM, VS, CDSM, CMM, DO, FI, PGV, BS, ER, SH, JHWD, TZ, MS, AV, EDL, SN, MC, LM, JH, CC, CAVM, KS, SR, LS, BA, VOS, MDS, SY, WASG, EC, JS, JVL, EL, PGDPL, KE, VC, JJAS, OKB, MMC and YA conceived and launched the EUSTAR database, collected data in their respective countries and offered critical comments regarding the manuscript. GR retrieved the death certificates of patients and of the general population and offered critical comments regarding the manuscript.

Funding This study was funded by the "Institut national de la santé et de la recherche médicale (INSERM)" (French National Health and Medical Research Institute).

Competing interests OD reports personal fees from 4D Science, grants and personal fees from Actelion, personal fees from Active Biotech, grants and personal fees from Bayer, personal fees from Biogen Idec, personal fees from BMS, grants and personal fees from Boehringer Ingelheim, personal fees from ChemomAb, personal fees from EpiPharm, personal fees from EspeRare Foundation, personal fees from Genentech/Roche, personal fees from GSK, personal fees from Inventiva, personal fees from Lilly, personal fees from Medac, personal fees from Mepha, personal fees from MedImmune, personal fees from Pharmacyclics, grants and personal fees from Pfizer, grants and personal fees from Sanofi, personal fees from Serodapharm, personal fees from Sinoxa, personal fees from AbbVie, personal fees from iQone Healthcare, outside the submitted work. In addition, OD has a patent mir-29 for the treatment of systemic sclerosis licensed. FI reports personal fees from AbbVie, personal fees from BMS, personal fees from MSD, personal fees from Novartis, outside the submitted work. JHWD reports personal fees from Actelion, grants and personal fees from Anamar, grants and personal fees from Bayer Pharma, grants and personal fees from Boehringer Ingelheim, grants from Celgene, personal fees from Galapagos, grants from GSK, grants and personal fees from Inventiva, personal fees from Pfizer, grants and personal fees from UCB, grants from Novartis, other from 4D Science, outside the submitted work. JvL reports personal fees from Pfizer, grants and personal fees from MSD, personal fees from Eli Lilly, personal fees from BMS, personal fees from Roche, outside the submitted work. Other coauthors have nothing to disclose.

Patient consent Obtained.

Ethics approval Local ethics committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data may be available from EUSTAR upon separate scientific request.

REFERENCES

1 Elhai M, Meune C, Avouac J, *et al*. Trends in mortality in patients with systemic sclerosis over 40 years: a systematic review and meta-analysis of cohort studies. *Rheumatology* 2012;51:1017–26.
2 Sampaio-Barros PD, Bortoluzzo AB, Marangoni RG, *et al*. Survival, causes of death, and prognostic factors in systemic sclerosis: analysis of 947 Brazilian patients. *J Rheumatol* 2012;39:1971–8.
3 Simeón-Aznar CP, Fonollosa-Plá V, Tolosa-Vilella C, *et al*. Registry of the Spanish Network for Systemic sclerosis: survival, Prognostic factors, and causes of death. *Medicine* 2015;94:e1728.
4 Ferri C, Sebastiani M, Lo Monaco A, *et al*. Systemic sclerosis evolution of disease pathomorphosis and survival. Our experience on Italian patients' population and review of the literature. *Autoimmun Rev* 2014;13:1026–34.
5 Rubio-Rivas M, Royo C, Simeón CP, *et al*. Mortality and survival in systemic sclerosis: systematic review and meta-analysis. *Semin Arthritis Rheum* 2014;44:208–19.
6 Komócsi A, Vorobcsuk A, Faludi R, *et al*. The impact of cardiopulmonary manifestations on the mortality of SSc: a systematic review and meta-analysis of observational studies. *Rheumatology* 2012;51:1027–36.
7 Bryan C, Knight C, Black CM, *et al*. Prediction of five-year survival following presentation with scleroderma: development of a simple model using three disease factors at first visit. *Arthritis Rheum* 1999;42:2660–5.
8 Fransen J, Popa-Diaconu D, Hesselstrand R, *et al*. Clinical prediction of 5-year survival in systemic sclerosis: validation of a simple prognostic model in EUSTAR centres. *Ann Rheum Dis* 2011;70:1788–92.
9 Beretta L, Santaniello A, Cappiello F, *et al*. Development of a five-year mortality model in systemic sclerosis patients by different analytical approaches. *Clin Exp Rheumatol* 2010;28:S18–27.
10 Scussel-Lonzetti L, Joyal F, Raynauld JP, *et al*. Predicting mortality in systemic sclerosis: analysis of a cohort of 309 French Canadian patients with emphasis on features at diagnosis as predictive factors for survival. *Medicine* 2002;81:154–67.
11 Zaridze D, Brennan P, Boreham J, *et al*. Alcohol and cause-specific mortality in Russia: a retrospective case-control study of 48,557 adult deaths. *Lancet* 2009;373:2201–14.
12 Kernéis S, Boëlle PY, Grais RF, *et al*. Mortality trends in systemic sclerosis in France and USA, 1980-1998: an age-period-cohort analysis. *Eur J Epidemiol* 2010;25:55–61.
13 Meier FM, Frommer KW, Dinser R, *et al*. Update on the profile of the EUSTAR cohort: an analysis of the EULAR Scleroderma Trials and Research group database. *Ann Rheum Dis* 2012;71:1355–60.
14 Tyndall AJ, Bannert B, Vonk M, *et al*. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Ann Rheum Dis* 2010;69:1809–15.
15 Israel RA, Rosenberg HM, Curtin LR. Analytical potential for multiple cause-of-death data. *Am J Epidemiol* 1986;124:161–79.
16 Avouac J, Amrouche F, Meune C, *et al*. Mortality profile of patients with rheumatoid arthritis in France and its change in 10 years. *Semin Arthritis Rheum* 2017;46:537–43.
17 Rey G, Aouba A, Pavillon G, *et al*. Cause-specific mortality time series analysis: a general method to detect and correct for abrupt data production changes. *Popul Health Metr* 2011;9:52.
18 Moreno-Betancur M, Sadaoui H, Piffaretti C, *et al*. Survival analysis with multiple causes of Death: extending the competing risks model. *Epidemiology* 2017;28:12–19.
19 Chiche L, Jourde-Chiche N, Bader-Meunier B, *et al*. Acute pancreatitis as a cause of mortality in pediatric systemic lupus erythematosus: Results of a multiple cause-of-death analysis in France. *Semin Arthritis Rheum* 2016;46:e6–e7.
20 van den Hoogen F, Khanna D, Fransen J, *et al*. Classification criteria for systemic sclerosis: an american college of rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis* 2013;2013:1747–55.
21 Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;81:515–26.
22 Steyerberg EW, Frank E, Jr H. *Regression modeling strategies with applications to linear models, logistic regression, and survival analysis*. . 2nd edn. Heidelberg: Springer, 2016:72. 1006–7.
23 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377–99.
24 White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med* 2009;28:1982–98.
25 Sandmeier B, Jäger VK, Nagy G, *et al*. Autopsy versus clinical findings in patients with systemic sclerosis in a case series from patients of the EUSTAR database. *Clin Exp Rheumatol* 2015;33:S75–9.
26 Allanore Y, Meune C. Primary myocardial involvement in systemic sclerosis: evidence for a microvascular origin. *Clin Exp Rheumatol* 2010;28:S48–53.

27 Allanore Y, Meune C, Vonk MC, *et al*. Prevalence and factors associated with left ventricular dysfunction in the EULAR Scleroderma Trial and Research group (EUSTAR) database of patients with systemic sclerosis. *Ann Rheum Dis* 2010;69:218–21.
28 Pène F, Hissem T, Bérezné A, *et al*. Outcome of patients with systemic sclerosis in the Intensive Care Unit. *J Rheumatol* 2015;42:1406–12.
29 Mouthon L, Mestre C, Bérezné A, *et al*. Low influenza vaccination rate among patients with systemic sclerosis. *Rheumatology* 2010;49:600–6.
30 Onishi A, Sugiyama D, Kumagai S, *et al*. Cancer incidence in systemic sclerosis: meta-analysis of population-based cohort studies. *Arthritis Rheum* 2013;65:1913–21.
31 Sears MR, Rea HH, de Boer G, *et al*. Accuracy of certification of deaths due to asthma. A national study. *Am J Epidemiol* 1986;124:1004–11.
32 Steyerberg EW, Harrell FE, Borsboom GJ, *et al*. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774–81.

CONCISE REPORT

Do depression and anxiety reduce the likelihood of remission in rheumatoid arthritis and psoriatic arthritis? Data from the prospective multicentre NOR-DMARD study

Brigitte Michelsen,[1,2,3] Eirik Klami Kristianslund,[1] Joseph Sexton,[1] Hilde Berner Hammer,[1] Karen Minde Fagerli,[1] Elisabeth Lie,[1] Ada Wierød,[4] Synøve Kalstad,[5] Erik Rødevand,[6] Frode Krøll,[7] Glenn Haugeberg,[3,8] Tore K Kvien[1]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211284).

[1]Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
[2]Department of Rheumatology, Hospital of Southern Norway Trust, Kristiansand, Norway
[3]Norwegian University of Science and Technology, Trondheim, Norway
[4]Department of Rheumatology, Vestre Viken/Drammen Hospital, Drammen, Norway
[5]Department of Rheumatology, University Hospital of Northern Norway, Tromsø, Norway
[6]Department of Rheumatology, St. Olavs Hospital, Trondheim, Norway
[7]Department of Rheumatology, Lillehammer Hospital for Rheumatic Diseases, Lillehammer, Norway
[8]Department of Rheumatology, Martina Hansens Hospital, Bærum, Norway

Correspondence to
Dr Brigitte Michelsen, Department of Rheumatology, Diakonhjemmet Hospital, Vinderen, N-0319 Oslo, Norway; brigitte_michelsen@yahoo.no

Received 9 February 2017
Revised 11 May 2017
Accepted 20 June 2017
Published Online First 21 July 2017

To cite: Michelsen B, Kristianslund EK, Sexton J, *et al*. *Ann Rheum Dis* 2017;**76**:1906–1910.

ABSTRACT

Objective To investigate the predictive value of baseline depression/anxiety on the likelihood of achieving joint remission in rheumatoid arthritis (RA) and psoriatic arthritis (PsA) as well as the associations between baseline depression/anxiety and the components of the remission criteria at follow-up.

Methods We included 1326 patients with RA and 728 patients with PsA from the prospective observational NOR-DMARD study starting first-time tumour necrosis factor inhibitors or methotrexate. The predictive value of depression/anxiety on remission was explored in prespecified logistic regression models and the associations between baseline depression/anxiety and the components of the remission criteria in prespecified multiple linear regression models.

Results Baseline depression/anxiety according to EuroQoL-5D-3L, Short Form-36 (SF-36) Mental Health subscale ≤56 and SF-36 Mental Component Summary ≤38 negatively predicted 28-joint Disease Activity Score <2.6, Simplified Disease Activity Index ≤3.3, Clinical Disease Activity Index ≤2.8, ACR/EULAR Boolean and Disease Activity Index for Psoriatic Arthritis ≤4 remission after 3 and 6 months treatment in RA (p≤0.008) and partly in PsA (p from 0.001 to 0.73). Baseline depression/anxiety was associated with increased patient's and evaluator's global assessment, tender joint count and joint pain in RA at follow-up, but not with swollen joint count and acute phase reactants.

Conclusion Depression and anxiety may reduce likelihood of joint remission based on composite scores in RA and PsA and should be taken into account in individual patients when making a shared decision on a treatment target.

INTRODUCTION

In today's treatment of rheumatoid arthritis (RA) and psoriatic arthritis (PsA), remission is the main target. However, far from all patients reach this target.[1]

Studies have suggested depression and anxiety to be associated with greater perception of pain in RA patients.[2 3] Furthermore, depression has been associated with reduced pain threshold and tolerance.[4] Depression is thought to promote sensitisation, to interfere with endogenous pain inhibition and to have a profound long-term influence on the shaping of pain responses and pain outcomes.[4] Depression and anxiety are more common in RA and PsA compared with the general population, with reported prevalence between 10%–42% in RA and 9%–37% in PsA.[5 6]

Only few studies have assessed the predictive value of depression and anxiety in RA suggesting poorer treatment outcomes.[3 7] Whether this can be confirmed in larger, prospective studies using various remission criteria, as well as in PsA, remains unexplored.

In this study, we aimed to investigate the predictive value of baseline depression and anxiety on the likelihood of achieving joint remission in patients with RA and PsA, in a large prospective multicentre observational register.[8] Secondary objectives were to explore the associations between baseline depression and anxiety and the components of the remission criteria at follow-up.

METHODS

Patients

We included RA and PsA patients from the Norwegian Disease-Modifying Anti-Rheumatic Drug (NOR-DMARD) register,[8] who started first tumour necrosis factor inhibitor (TNFi) with or without comedication with methotrexate (MTX), or started MTX as first DMARD, all patients included only once (see online supplementary figure S1). NOR-DMARD is a prospective observational multicentre study initiated in December 2000 including patients >18 years with inflammatory arthropathies starting biological or synthetic DMARDs. Assessments were made at baseline, 3, 6 and 12 months and then yearly. The study comprises five centres in different geographical regions in Norway.

The patients in the current analyses were included between 1 March 2006 (introduction of EuroQoL-5D-3L (EQ-5D-3L)[9] and 6 November 2012 and followed until 15 April 2013. Written informed consent was obtained from each patient. The study was approved by the National Data Inspectorate and by the Regional Ethics Committee.

Table 1 Demographics and baseline disease activity measures for patients with versus without baseline depression/anxiety according to the EQ-5D-3L criterion

	Rheumatoid arthritis (n=1326)			Psoriatic arthritis (n=728)		
	Depressed/anxious (n=573)	Not depressed/anxious (n=753)	p Value	Depressed/anxious (n=326)	Not depressed/anxious (n=402)	p Value
Age (years), mean (SD)	53.5 (13.2)	55.3 (13.8)	0.02	46.9 (12.6)	48.9 (12.4)	0.03
Female, %	75.1	62.3	0.01	59.7	44.1	<0.001
Disease duration (years), median (25th, 75th percentile)	0.27 (0.006, 5.3)	0.27 (0.003, 5.0)	0.72	0.95 (0.08, 6.4)	0.70 (0.04, 7.1)	0.78
Currently smoking n/total (%)	185/572 (32.3)	192/744 (25.8)	0.009	109/323 (33.7)	298/398 (25.1)	0.01
Joint pain, mean (SD)	51.8 (24.9)	40.5 (24.9)	<0.001	52.6 (22.4)	42.2 (22.9)	<0.001
32 tender joints, median (25th, 75th percentile)	9 (4, 14)	6 (3, 11)	<0.001	5 (2, 10)	4 (2, 8)	0.002
32 swollen joints, median (25th, 75th percentile)	6 (3, 10)	5 (3, 10)	0.26	2 (1, 5)	3 (1, 6)	0.16
EGA, mean (SD)	37.6 (19.0)	35.4 (19.3)	0.04	33.8 (16.3)	31.1 (15.7)	0.03
PGA, mean (SD)	55.2 (24.3)	42.5 (24.1)	<0.001	56.2 (22.1)	44.6 (23.0)	0.001
DAS28ESR, mean (SD)	4.8 (1.3)	4.5 (1.4)	<0.001	4.1 (1.3)	3.8 (1.2)	0.002
DAS28CRP, mean (SD)	4.6 (1.2)	4.3 (1.3)	<0.001	4.0 (1.1)	3.7 (1.1)	0.002
SDAI (25th, 75th percentile)	22.4 (15.6, 32.3)	19.2 (12.1, 30.5)	<0.001	16.5 (11.4, 23.3)	13.9 (9.6, 21.0)	0.001
CDAI (25th, 75th percentile)	21.0 (14.2, 30.4)	17.4 (10.9, 27.5)	<0.001	14.8 (10.6, 21.5)	12.8 (8.7, 18.9)	0.001
MHAQ (25th, 75th percentile)	0.8 (0.4, 1.1)	0.5 (0.1, 0.9)	<0.001	0.6 (0.4, 1.0)	0.5 (0.1, 0.8)	<0.001
Modified DAPSA, median (25th, 75th percentile)	27.4 (19.6, 38.0)	22.2 (14.5, 34.1)	<0.001	21.1 (14.3, 28.2)	17.5 (11.7, 24.7)	<0.001
ESR, mm/h, median (25th, 75th percentile)	19 (11, 35)	21 (11, 37)	0.19	17 (7, 30)	15 (8, 26)	0.42
CRP, mg/L, median (25th, 75th percentile)	8 (3, 21)	8 (5, 23)	0.29	7 (3, 20)	6 (4, 155)	0.70
Erosive disease at baseline, n/total (%)	249/558 (44.6)	317/731 (43.4)	0.65	115/320 (35.9)	144/385 (37.4)	0.69

CDAI, Clinical Disease Activity Index; CRP, C reactive protein; DAPSA, Disease Activity Index for Psoriatic Arthritis; DAS28ESR, 28-joint Disease Activity Score; EGA, evaluator's global assessment on a 0–100 Visual Analogue Scale (VAS) scale; ESR, erythrocyte sedimentation rate; MHAQ, Modified Health Assessment Questionnaire; PGA, patient's global assessment on a 0–100 VAS scale; SDAI, Simplified Disease Activity Index.

Assessments

Assessments included 32 tender/swollen joint count (28 joint count with addition of ankles (0–2) and metatarso-phalangeal joints (0–2)), visual analogue scales (0–100 mm) for joint pain, patient's and evaluator's global assessments, Modified Health Assessment Questionnaire,[10] erythrocyte sedimentation rate (mm/hour), C reactive protein (CRP; mg/L), disease duration, current smoking status (yes/no) and erosive disease at baseline (yes/no). Twenty-eight-joint Disease Activity Score (DAS28),[11] Simplified Disease Activity Index (SDAI)[11] and Clinical Disease Activity Index (CDAI)[11] were computed, as well as a modified Disease Activity Index for Psoriatic Arthritis (DAPSA)[12] using 32 instead of 66/68 joint count and joint pain instead of pain. The following criteria for depression/anxiety were applied: (1) EQ-5D-3L question 5: 'I am not or moderately/extremely anxious or depressed',[9] (2) Medical Outcomes Survey Short Form-36 (SF-36) Mental Health subscale (SF-36MH)≤56[13] and (3) SF-36 Mental Component Summary (SF-36MCS)≤38.[13] The following remission criteria were explored: (1) DAS28 <2.6,[11] (2) SDAI ≤3.3,[11 14] (3) CDAI ≤2.8,[11] (4) ACR/EULAR Boolean[11] and (5) DAPSA ≤4.[12]

Statistics

For the patients' demographic and baseline characteristics medians (25th and 75th percentiles) were calculated for non-normally and means (SD) for normally distributed data. Continuous measures were compared using Mann-Whitney U-test or independent t-test as appropriate. Proportions were analysed using X^2 test.

In the main analyses, as well as in subgroup analyses of TNFi/MTX-treated patients, the predictive value of depression/anxiety for DAS28, SDAI, CDAI, ACR/EULAR Boolean and DAPSA remission after 3 and 6 months treatment was explored in prespecified logistic regression models adjusted for age, sex, disease duration and smoking.

The analyses were performed as completer analyses and without adjustment for multiple comparisons. Sensitivity analyses with additional adjustment for baseline erosive disease and CRP were performed.

In secondary analyses, the components of the remission criteria at 3 and 6 months follow-up were explored for patients with and without baseline depression/anxiety in prespecified multiple linear regression models adjusted for age, sex, disease duration and smoking, as well as with and without adjustment for baseline values of the explored outcome variables. Statistical tests were performed using SPSS for Windows V.21.0.

RESULTS

Demographics and baseline disease activity measures according to EQ-5D-3L depression/anxiety status are shown in table 1. Number/total (%) of depressed/anxious patients according to SF-36MH≤56/ SF-36MCS≤38 was 238/1311 (18.2)/ 381/1277 (29.8) in RA and 115/724 (15.9)/ 182/708 (25.7) in PsA, respectively.

Associations between baseline depression/anxiety and achievement of joint remission

The proportion of patients achieving joint remission was overall lower in patient with versus without baseline depression/anxiety (figure 1).

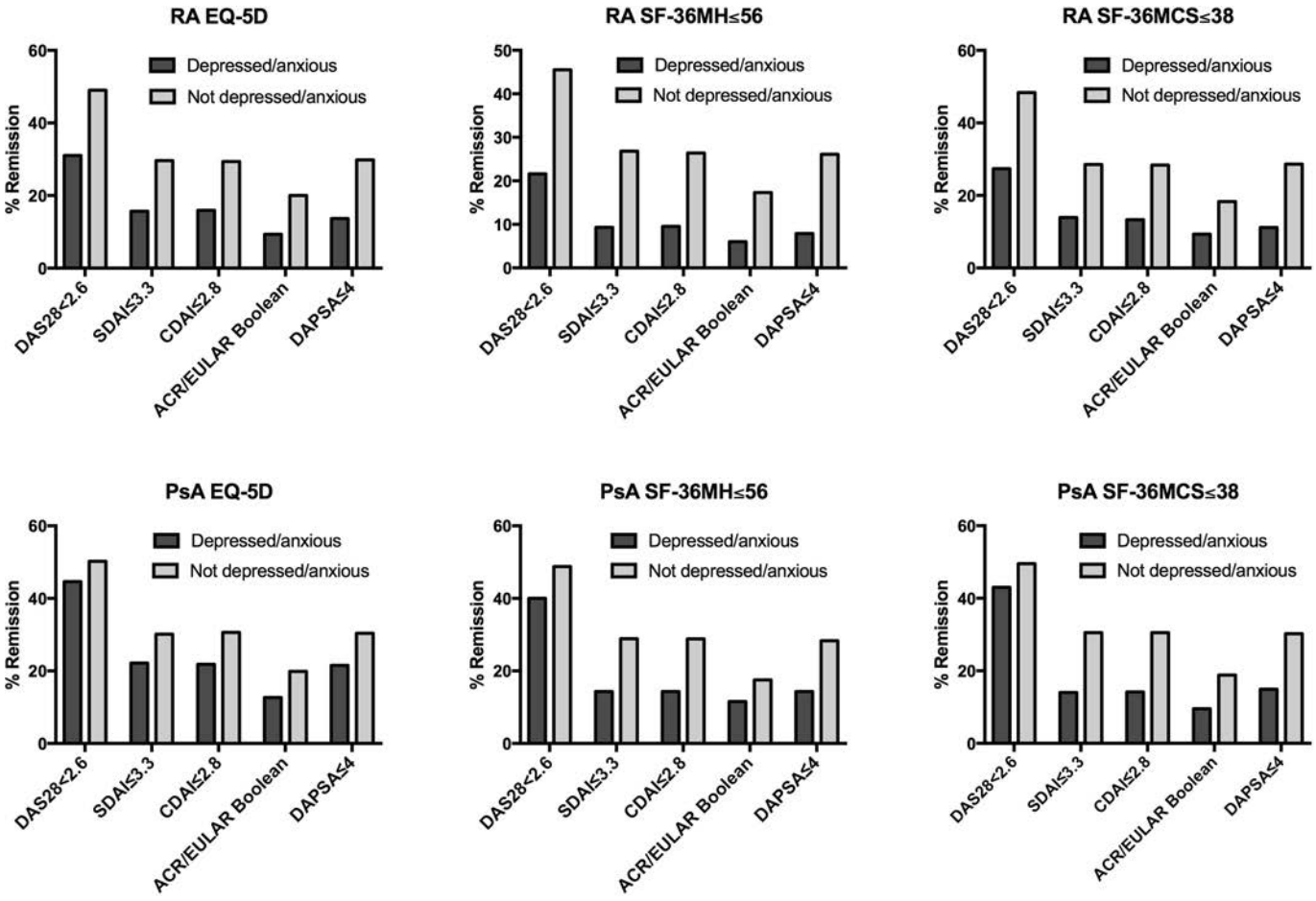


Figure 1 Bar charts of percentages of rheumatoid arthritis and psoriatic arthritis patients in remission at 6 months across presence versus absence of depression/anxiety. CDAI, Clinical Disease Activity Index; DAPSA, Disease Activity Index for Psoriatic Arthritis; DAS28, 28-joint Disease Activity Score; EQ-5D-3L, question 5 from the EuroQol-5D-3L questionnaire; PsA, psoriatic arthritis; RA, rheumatoid arthritis; SDAI, Simplified Disease Activity Index; SF-36MH, Medical Outcomes Survey Short Form-36 Mental Health subscale; SF-36MCS, Medical Outcomes Survey Short Form-36 Mental Component Summary.

Baseline depression/anxiety according to EQ-5D-3L, SF-36MH≤56 and SF-36MCS≤38 negatively predicted achievement of DAS28, SDAI, CDAI, ACR/EULAR Boolean and DAPSA remission after 3 and 6 months treatment in RA. Corresponding findings in PsA also showed consistently lower point estimates but did not reach significance for all the analyses (table 2).

The findings were consistent also in sensitivity analyses (see online supplementary table S1). In subgroup analyses of TNFi/MTX-treated patients, baseline depression/anxiety according to EQ-5D-3L, SF36-MH≤56 and SF36-MCS≤38 were found to be negative predictors of remission at 6 months in RA, but only partly in PsA (see online supplementary tables S2–4).

Secondary outcomes analyses

Baseline depression/anxiety was associated with increased patient's and evaluator's global assessment, joint pain and tender joint count at 3 and 6 months in RA and with increased patient's global assessment and joint pain in PsA, but not with swollen joint count or levels of acute phase reactants (see online supplementary tables S5a-c and S6a-c). Additional adjustment for baseline values of the explored outcome variables did not change the main results.

DISCUSSION

In this prospective multicentre observational study, baseline depression and anxiety were found to be strong negative predictors of remission after 3 and 6 months treatment in RA, and partly in PsA, using joint-specific measures of remission. The findings were valid across different depression/anxiety and remission criteria, as well as in subgroup analyses of TNFi-treated and MTX-treated patients.

The thresholds SF-36MH≤56 and SF-36MCS≤38 were recently found to have good sensitivity and specificity to detect depression and anxiety in RA patients.[13] The EULAR recommendation to use one single question to screen for depression in chronic inflammatory rheumatic diseases is in accordance with the EQ-5D-3L criterion,[15] which may measure slightly different aspects compared with the more comprehensive SF-36 derived criteria.

The prespecified logistic regression model was adjusted for age, sex, disease duration and smoking, but not for recently identified predictors in NOR-DMARD for only one or some of the explored remission criteria, nor for patient-reported outcomes (PROs) or composite scores comprising PROs, as these

Table 2 The predictive value of baseline depression/anxiety for achievement of joint remission after 3 and 6 months treatment (logistic regression analyses)

Months		DAS28ESR<2.6	SDAI≤3.3	CDAI≤2.8	ACR/EULAR Boolean	DAPSA≤4
OR (95% CI) for joint remission in patients with versus without baseline depression/anxiety according to the EQ-5D-3L criterion						
3	RA	0.66 (0.49 to 0.90), p=0.008 (n=855)	0.57 (0.39 to 0.83), p=0.003 (n=924)	0.49 (0.34 to 0.70), p<0.001 (n=991)	0.46 (0.30 to 0.71), p<0.001 (n=1017)	0.58 (0.39 to 0.86), p=0.006 (n=924)
	PsA	0.74 (0.48 to 1.12), p=0.15 (n=416)	0.59 (0.37 to 0.96), p=0.03 (n=455)	0.50 (0.32 to 0.78), p=0.002 (n=522)	0.51 (0.30 to 0.88), p=0.02 (n=533)	0.54 (0.33 to 0.91), p=0.02 (n=455)
6	RA	0.48 (0.35 to 0.66), p<0.001 (n=715)	0.45 (0.31 to 0.65), p<0.001 (n=788)	0.46 (0.33 to 0.65), p<0.001 (n=880)	0.44 (0.29 to 0.66), p<0.001 (n=886)	0.37 (0.25 to 0.54), p<0.001 (n=797)
	PsA	0.92 (0.59 to 1.44), p=0.73 (n=371)	0.70 (0.44 to 1.12), p=0.14 (n=426)	0.69 (0.45 to 1.06), p=0.09 (n=488)	0.61 (0.37 to 1.01), p=0.05 (n=492)	0.69 (0.43 to 1.10), p=0.12 (n=428)
OR (95% CI) for joint remission in patients with versus without baseline depression/anxiety according to the SF-36MH≤56 criterion						
3	RA	0.48 (0.31 to 0.73), p=0.001 (n=841)	0.37 (0.20 to 0.67), p=0.001 (n=911)	0.35 (0.19 to 0.61), p<0.001 (n=977)	0.22 (0.09 to 0.50), p<0.001 (n=1003)	0.15 (0.06 to 0.34), p<0.001 (n=924)
	PsA	0.59 (0.33 to 1.05), p=0.07 (n=414)	0.36 (0.16 to 0.80), p=0.01 (n=453)	0.49 (0.25 to 0.99), p=0.046 (n=520)	0.59 (0.27 to 1.31), p=0.20 (n=530)	0.37 (0.17 to 0.82), p=0.01 (n=463)
6	RA	0.33 (0.21 to 0.53), p<0.001 (n=709)	0.28 (0.16 to 0.52), p<0.001 (n=782)	0.30 (0.17 to 0.54), p<0.001 (n=874)	0.30 (0.50 to 0.61), p=0.001 (n=878)	0.25 (0.13 to 0.47), p<0.001 (n=791)
	PsA	0.78 (0.41 to 1.46), p=0.43 (n=368)	0.40 (0.18 to 0.86), p=0.02 (n=423)	0.41 (0.20 to 0.86), p=0.02 (n=485)	0.65 (0.29 to 1.43), p=0.28 (n=488)	0.43 (0.20 to 0.91), p=0.03 (n=424)
OR (95% CI) for joint remission in patients with versus without baseline depression/anxiety according to the SF-36MCS≤38 criterion						
3	RA	0.67 (0.48 to 0.93), p=0.02 (n=819)	0.44 (0.28 to 0.69), p<0.001 (n=890)	0.40 (0.26 to 0.62), p<0.001 (n=954)	0.26 (0.14 to 0.47), p<0.001 (n=979)	0.33 (0.20 to 0.53), p<0.001 (n=903)
	PsA	0.60 (0.36 to 0.99), p=0.04 (n=403)	0.35 (0.18 to 0.68), p=0.002 (n=443)	0.44 (0.24 to 0.78), p=0.005 (n=505)	0.43 (0.21 to 0.87), p=0.02 (n=517)	0.39 (0.21 to 0.74), p=0.004 (n=454)
6	RA	0.40 (0.28 to 0.57), p<0.001 (n=693)	0.40 (0.26 to 0.61), p<0.001 (n=765)	0.39 (0.26 to 0.58), p<0.001 (n=854)	0.45 (0.28 to 0.72), p=0.001 (n=859)	0.30 (0.19 to 0.48), p<0.001 (n=774)
	PsA	0.76 (0.44 to 1.32), p=0.32 (n=356)	0.39 (0.20 to 0.75), p=0.005 (n=411)	0.43 (0.23 to 0.79), p=0.006 (n=471)	0.55 (0.28 to 1.09), p=0.09 (n=474)	0.40 (0.21 to 0.78), p=0.007 (n=412)

Analyses were adjusted for age, gender, disease duration and smoking.
CDAI, Clinical Disease Activity Index; DAPSA, Disease Activity Index for Psoriatic Arthritis; DAS28, 28-joint disease activity score; EQ-5D-3L, question five from the EuroQol-5D-3L questionnaire; PsA, psoriatic arthritis (n=728); RA, rheumatoid arthritis (n=1326); SDAI, Simplified Disease Activity Index; SF-36MH, Medical Outcomes Survey Short Form-36 Mental Health subscale; SF-36MCS, Medical Outcomes Survey Short Form-36 Mental Component Summary.

may be symptoms of/ influenced by depression and anxiety.[2 4 16] The validity of the findings reported in table 2 was confirmed in sensitivity analyses with adjustment for additional covariates (see online supplementary table S1).

This study is in line with a recent study in which self-reported lifetime history of depression was associated with decreased likelihood of 6 months CDAI remission in RA.[7] However, associations between baseline depression and the components of CDAI at follow-up were not confirmed. Furthermore, in post hoc analyses of clinical trial data in RA. persisting depression/anxiety assessed by EQ-5D was associated with a decreased likelihood of DAS28 remission as well as higher patient's global assessment and tender joint count after 2 years.[3] By contrast, in a smaller study in RA, depression/anxiety did not reach significance as predictors of 1-year DAS28 remission.[17] Still, baseline depression/anxiety was associated with higher tender joint count and patient's global assessment at follow-up.[17]

We have recently reported reduced likelihood of remission among RA and PsA patients with discordance between baseline patient's and evaluator's global assessment as well as tender and swollen joint count.[18] Discordance between more subjectively and more objectively weighted measures of disease activity may be reflected by psychosocial factors like depression.[4]

Major limitations of the study are lack of consensus on validated remission criteria in PsA, lack of established diagnoses of depression and anxiety, as well as lack of psoriasis severity information, as psoriasis may have major implications for mental health. Furthermore, the assessment of 32 and not 66/68 joint count may have led to overestimation of remission rates in PsA and missing data handled by completer analyses may have affected the generalisability of the results. Finally, the modification of DAPSA with 32 instead of 66/68 joint count and joint pain instead of pain is not validated for PsA and DAPSA is developed as remission criterion for PsA and not RA.[12]

The major strength of this study is the prospective observational multicentre design and the inclusion of large cohorts of RA and PsA patients over a long time span. Furthermore, this is the first study to assess the predictive value of depression and anxiety on achievement of remission in PsA and the first study using EQ-5D as well as SF-36-derived depression/anxiety criteria as predictors of remission in RA. The consistent findings for these different predictors support the validity and the robustness of the results.

Depression and anxiety are frequently occurring disorders but are not widely considered in routine care of arthritis patients, although of considerable economic impact to the society.[6 19] Increased emphasis on the negative predictive value of depression and anxiety ought to be considered as part of a treat-to-target strategy in patients not reaching remission. In particular, alternative targets to composite scores may be considered in the shared decision making between the patient and healthcare provider. Thus, depression and anxiety should be taken into account according to recommendation number 5 (RA) and 6 (PsA) in the treat-to-target recommendations.[20 21]

In conclusion, depression and anxiety were found to be strong, negative predictors of joint remission at 3 and 6 months treatment in RA and partly in PsA, according to various remission as well as depression/anxiety criteria. Depression and anxiety were associated to more subjectively weighted measures, but not acute phase reactants and swollen joint count during follow-up. These observations support a focus on depression and anxiety as comorbidities in a treat-to-target strategy.

Acknowledgements The authors would like to thank the patients for participating in this study and the local rheumatology staff for data collection.

Contributors BM, EKK, JS, HBH, KMF, EL, GH and TKK were responsible for study design. AW, SK, ER, FK and TKK were responsible for data acquisition. BM analysed the data and wrote the manuscript. All authors critically revised the manuscript and approved the final version.

Funding The study was funded through a clinical research fellowship from Diakonhjemmet Hospital, originating from South-Eastern Health Authority. Data collection in NOR-DMARD was partly funded through unrestricted grants from Abbvie, BMS, MSD, Pfizer (Wyeth), Roche and UCB.

Competing interests HBH has received fees for speaking and/or consulting from AbbVie, Pfizer, UCB, Roche, MSD, BMS and Novartis. TKK has received fees for speaking and/or consulting from AbbVie, Biogen, BMS, Boehringer Ingelheim, Celgene, Celltrion, Eli Lilly, Epirus, Hospira, Merck-Serono, MSD, Mundipharma, Novartis, Oktal, Orion Pharma, Hospira/Pfizer, Roche, Sandoz and UCB and received research funding to Diakonhjemmet Hospital from AbbVie, BMS, MSD, Pfizer, Roche and UCB. All other authors have declared that no competing interests exist.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Aga AB, Lie E, Uhlig T, *et al*. Time trends in disease activity, response and remission rates in rheumatoid arthritis during the past decade: results from the NOR-DMARD study 2000-2010. *Ann Rheum Dis* 2015;74:381–8.
2 Boyden SD, Hossain IN, Wohlfahrt A, *et al*. Non-inflammatory causes of pain in patients with Rheumatoid Arthritis. *Curr Rheumatol Rep* 2016;18:30.
3 Matcham F, Norton S, Scott DL, *et al*. Symptoms of depression and anxiety predict treatment response and long-term physical health outcomes in rheumatoid arthritis: secondary analysis of a randomized controlled trial. *Rheumatology* 2016;55:268–78.
4 Edwards RR, Cahalan C, Calahan C, *et al*. Pain, catastrophizing, and depression in the rheumatic diseases. *Nat Rev Rheumatol* 2011;7:216–24.
5 McDonough E, Ayearst R, Eder L, *et al*. Depression and anxiety in psoriatic disease: prevalence and associated factors. *J Rheumatol* 2014;41:887–96.
6 Matcham F, Rayner L, Steer S, *et al*. The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis. *Rheumatology* 2013;52:2136–48.
7 Rathbun AM, Harrold LR, Reed GW. A prospective evaluation of the effects of prevalent depressive symptoms on disease activity in rheumatoid Arthritis Patients Treated with biologic response modifiers. *Clin Ther* 2016;38:1759–72.
8 Kvien TK, Lie E, *et al*. A Norwegian DMARD register: prescriptions of DMARDs and biological agents to patients with inflammatory rheumatic diseases. *Clin Exp Rheumatol* 2005;23:S188–94.
9 EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
10 Pincus T, Summey JA, Soraci SA, *et al*. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983;26:1346–53.
11 Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet* 2016;388:2023–38.
12 Schoels MM, Aletaha D, Alasti F, *et al*. Disease activity in psoriatic arthritis (PsA): defining remission and treatment success using the DAPSA score. *Ann Rheum Dis* 2016;75:811–8.
13 Matcham F, Norton S, Steer S, *et al*. Usefulness of the SF-36 Health Survey in screening for depressive and anxiety disorders in rheumatoid arthritis. *BMC Musculoskelet Disord* 2016;17:224.
14 Salaffi F, Ciapetti A, Carotti M, *et al*. Disease activity in psoriatic arthritis: comparison of the discriminative capacity and construct validity of six composite indices in a real world. *Biomed Res Int* 2014;2014:1–12.
15 Baillet A, Gossec L, Carmona L, *et al*. Points to consider for reporting, screening for and preventing selected comorbidities in chronic inflammatory rheumatic diseases in daily practice: a EULAR initiative. *Ann Rheum Dis* 2016;75:965–73.
16 Uhlig T, Lie E, Norvang V, *et al*. Achievement of remission and low disease activity definitions in patients with rheumatoid Arthritis in clinical practice: results from the NOR-DMARD Study. *J Rheumatol* 2016;43:716–23.
17 Matcham F, Ali S, Irving K, *et al*. Are depression and anxiety associated with disease activity in rheumatoid arthritis? A prospective study. *BMC Musculoskelet Disord* 2016;17:155.
18 Michelsen B, Kristianslund EK, Hammer HB, *et al*. Discordance between tender and swollen joint count as well as patient's and evaluator's global assessment may reduce likelihood of remission in patients with rheumatoid arthritis and psoriatic arthritis: data from the prospective multicentre NOR-DMARD study. *Ann Rheum Dis* 2016;76:708–11.
19 Uhlig T, Moe RH, Kvien TK. The burden of disease in rheumatoid arthritis. *Pharmacoeconomics* 2014;32:841–51.
20 Smolen JS, Breedveld FC, Burmester GR, *et al*. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;75:3–15.
21 Smolen JS, Braun J, Dougados M, *et al*. Treating spondyloarthritis, including ankylosing spondylitis and psoriatic arthritis, to target: recommendations of an international task force. *Ann Rheum Dis* 2014;73:6–16.

CONCISE REPORT

Role of erosions typical of rheumatoid arthritis in the 2010 ACR/EULAR rheumatoid arthritis classification criteria: results from a very early arthritis cohort

Gina Hetland Brinkmann,[1,2] Ellen S Norli,[2,3] Pernille Bøyesen,[2] Désirée van der Heijde,[4] Lars Grøvle,[1] Anne J Haugen,[1] Halvor Nygaard,[5] Olav Bjørneboe,[3] Cathrine Thunem,[6] Tore K Kvien,[2] Maria D Mjaavatten,[2] Elisabeth Lie[2]

Handling editor Gerd R Burmester

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211350).

[1]Department of Rheumatology, Østfold Hospital Truststfold Hospital Trust, Grålum, Norway
[2]Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
[3]Department of Rheumatology, Martina Hansen Hospital, Bærum, Norway
[4]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[5]Department of Rheumatology, Lillehammer Hospital for Rheumatic Diseases, Lillehammer, Norway
[6]Department of Rheumatology, Betanien Hospital, Skien, Norway

Correspondence to
Dr Gina Hetland Brinkmann, Department of Rheumatology, Østfold Hospital Trust, Postbox 300, Grålum, N-1714, Norway; gina_hetland@hotmail.com

Received 21 February 2017
Revised 31 May 2017
Accepted 5 July 2017
Published Online First 9 August 2017

To cite: Brinkmann GH, Norli ES, Bøyesen P, *et al*. *Ann Rheum Dis* 2017;**76**:1911–1914.

ABSTRACT

Objective To determine how the European League Against Rheumatism (EULAR) definition of erosive disease (erosion criterion) contributes to the number of patients classified as rheumatoid arthritis (RA) according to the 2010 American College of Rheumatology/EULAR RA classification criteria (2010 RA criteria) in an early arthritis cohort.

Methods Patients from the observational study Norwegian Very Early Arthritis Clinic with joint swelling ≤16 weeks, a clinical diagnosis of RA or undifferentiated arthritis, and radiographs of hands and feet were included. Erosive disease was defined according to the EULAR definition accompanying the 2010 RA criteria. We calculated the additional number of patients being classified as RA based on the erosion criteria at baseline and during follow-up.

Results Of the 289 included patients, 120 (41.5%) fulfilled the 2010 RA criteria, whereas 15 (5.2%) fulfilled only the erosion criterion at baseline. 118 patients had radiographic follow-up at 2 years, of whom 6.8% fulfilled the 2010 RA criteria and only one patient fulfilled solely the erosion criterion during follow-up.

Conclusion Few patients with early arthritis were classified as RA based on solely the erosion criteria, and of those who did almost all did so at baseline.

INTRODUCTION

The aim with the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) rheumatoid arthritis (RA) classification criteria (2010 RA criteria) was to identify patients at high risk for developing persistent erosive and/or inflammatory disease in the early stage of the disease. Erosions were not considered for inclusion in the scoring system by the ACR/EULAR working group because patients with erosions typical of RA were presumed to have prima facie evidence of RA.[1] Later, a task force suggested that the presence of ≥3 joints with typical erosions is sufficient to classify patients as having RA based on erosive disease on radiographs alone (erosion criterion).[2 3]

The main objective of this study was to assess to what degree the EULAR definition of erosive disease contributes to the number of patients classified as RA according to the 2010 RA criteria.

METHODS

Setting and patient selection

The current analyses were based on data from the observational, prospective Norwegian Very Early Arthritis Clinic study, including patients with ≥1 clinically swollen joint of ≤16 weeks' duration.[4] The cohort included 1118 patients (age 18–75 years) between years 2004 and 2010 with study visits at baseline and after 3, 6, 12 and 24 months. Patients with joint swelling due to trauma, osteoarthritis, crystal arthritis or septic arthritis were excluded. The study was approved by the Regional Ethics Committee of Southern Norway.

In this current study patients with a clinical diagnosis other than RA or undifferentiated arthritis (UA) at baseline were excluded. We included the remaining patients with radiographs of hands and feet at baseline (online supplementary file 1). A subset of these patients had radiographs both at baseline and at 2 years of follow-up (online supplementary file 2, figure S2). The 2010 RA criteria were retrospectively applied at baseline and cumulatively at follow-up visits.

Data collection and radiographic assessment

The full data collection has been described elsewhere.[4] In the current study, conventional hand and feet radiographs were performed at baseline and at 24 months in patients included from year 2007 and onwards.

A trained reader, blinded to patient characteristics, scored radiographs of hands and feet according to the van der Heijde modified Sharp score method.[5] The time order of the radiographs was known. We defined erosive disease according to the EULAR definition accompanying the 2010 RA criteria: 'Erosive disease for use in the 2010 RA criteria is defined when an erosion (defined as a cortical break) is seen at at least three separate joints at any of the following sites: the proximal interphalangeal (PIP) joints, metacarpophalangeal (MCP) joints, the wrist (counted as one joint) and the metatarsophalangeal (MTP) joints on radiographs of both hands and feet'.[3] The wrist includes the carpometacarpal(CMC) bone, trapezium, scaphoid, lunate, radial and ulnar bone. The 1st interphalangeal (IP1) joint of the feet is included in the MTP joints.

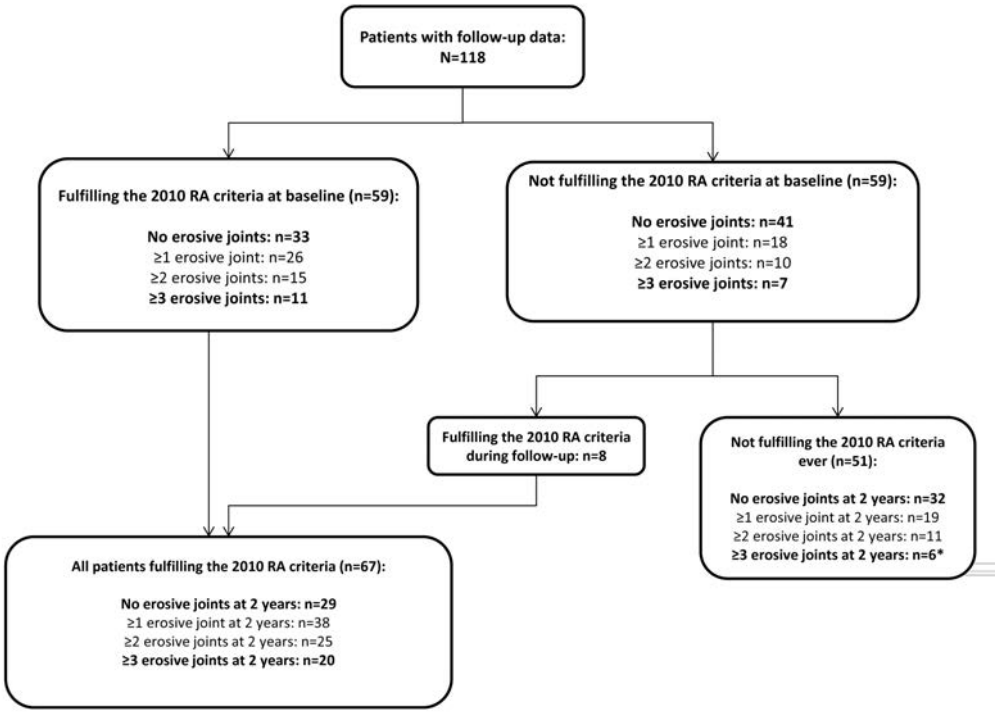


Figure 1 Fulfilment of the 2010 RA criteria and presence of erosive joints in patients at baseline and during follow-up. *Five patients had ≥3 erosive joints at baseline and one patient developed ≥3 erosions during follow-up. RA, rheumatoid arthritis.

Analyses

The number of patients being classified as RA based on the 2010 RA criteria and on the erosion criterion (without fulfilling the 2010 RA criteria), as well as the number of patients with ≥1 and ≥2 erosive joints, were calculated at baseline and during follow-up. We also evaluated the distribution of erosive joints. Additionally, incident erosions during follow-up in patients without baseline erosions were determined.

RESULTS

Fulfilment of the 2010 RA criteria and the EULAR definition of erosive disease (erosion criterion)

The baseline characteristics of the 289 patients included in the current study are shown in online supplementary file, table S1. One hundred and twenty (41.5%) patients fulfilled the 2010 RA criteria at baseline, of whom 49 (40.8%) patients had ≥1 erosive joint and 17 (14.2%) fulfilled the erosion criterion. Of the remaining 169 patients not fulfilling the 2010 RA criteria, 15 patients fulfilled the erosion criterion, 27 had ≥2 erosive joints and 55 patients had ≥1 erosive joint.

One hundred and eighteen patients had 2-year radiographic follow-up data. Presence/absence of erosions at baseline and follow-up in relation to the fulfilment of the 2010 RA criteria is shown in figure 1. Of all 118 patients, 8 additional patients (6.8%) fulfilled the 2010 RA criteria during follow-up, while 6 fulfilled the erosion criterion alone, of whom 5 patients at baseline.

Characteristics of patients with erosions

Table 1 presents the baseline characteristics of patients not fulfilling and patients fulfilling the 2010 RA criteria. The 15 patients who fulfilled the erosion criterion but not the 2010 RA criteria were all seronegative. They also had numerically shorter duration of joint swelling, were more often men and had fewer involved joints than those who fulfilled the 2010 RA criteria. One of these 15 patients received the final clinical diagnosis RA, while the rest of the patients were diagnosed with UA (n=10), reactive arthritis (n=1), chondrocalcinosis (n=1) or osteoarthritis (n=2). Furthermore, 2 of the 15 patients were treated with disease modifying antirheumatic drugs (DMARDs) (1 patient with RA and 1 patient with reactive arthritis).

Distribution of erosive joints at baseline in patients not fulfilling the 2010 RA criteria and incident erosions during follow-up

As shown in table 2, the MTP and PIP joints were the most frequently affected joints at baseline. Of the 169 patients not fulfilling the 2010 RA criteria, 40 patients had ≥1 hand erosion, 28 had ≥1 foot erosion and 13 had erosions in both hands and feet at baseline. Among patients with no baseline erosions (n=74), 13 (17.6%) developed erosions during follow-up (PIP joints n=3, MCP n=4, wrist n=4, MTP joints n=12), of whom 7 fulfilled the 2010 RA criteria at baseline, and 2 patients fulfilled both the 2010 RA criteria (at baseline) and the erosion criterion (during follow-up).

DISCUSSION

Nearly 42% of the 289 patients fulfilled the 2010 RA criteria at baseline, and only additional 6.8% (eight patients) did so during the 2-year follow-up. In total, 16 patients were classified as RA based on the erosion criterion (15 at baseline and 1 during follow-up). We obtained nearly the same percentage fulfilling the erosion criterion at baseline as observed in the two cohorts used to define the erosion criterion (5.2% vs 3.3%, respectively).[3]

Finding patients with ≥3 erosive joints at baseline is surprising considering the short duration of joint swelling, that is, ≤16

Table 1 Baseline characteristics of patients with and without erosions not fulfilling and patients fulfilling the 2010 RA criteria

	Patients not fulfilling the 2010 RA criteria at baseline (n=169)				Patients fulfilling the 2010 RA criteria at baseline (n=120)
	No erosions (n=114)	≥1 erosion (n=55)	≥2 erosions (n=27)	≥3 erosions* (n=15)	
Age, mean (SD)	41.4 (14.1)	56.4 (11.2)	59.6 (10.0)	61.8 (10.9)	52.3 (13.7)
Female gender, n (%)	59 (51.8)	26 (47.3)	13 (48.1)	6 (40.0)	72 (60.0)
Duration of joint swelling (days), median (25, 75 percentile)	31 (11, 67)	36 (9, 67)	55 (17, 76)	48 (17, 76)	67 (37, 88)
Ever smoker, n (%)	54 (47.4)	33 (60.0)	18 (66.7)	11 (73.3)	89 (74.2)
RF and/or ACPA positive, n (%)	6 (5.4)	6 (10.9)	2 (7.4)	0	78 (65.5)
– RF (IgM or IgA) positive, n (%)	5 (4.5)	2 (3.6)	0	0	61 (51.3)
– ACPA positive, n (%)	3 (2.7)	5 (9.1)	2 (7.4)	0	67 (56.3)
ESR, mm/hour, median (25, 75 percentile)	20 (9, 36.25)	20 (10, 37)	32.5 (14, 41)	34 (14, 44)	28 (14.25, 44.75)
CRP, mg/L, median (25, 75 percentile)	12 (2, 41.75)	10 (3, 30)	10 (4, 30)	16 (6, 28)	15.5 (6, 38)
68-SJC, median (25, 75 percentile)	2 (1, 3)	2 (1, 4)	2 (1, 5)	3 (2, 6)	11 (5, 17)
28-TJC, median (25, 75 percentile)	1 (1, 2)	2 (1, 4)	1 (1, 4)	1 (1, 3)	7.5 (4, 12.75)
Small joint involvement, n (%)	51 (44.7)	35 (63.6)	18 (66.7)	13 (86.7)	118 (98.3)
Polyarticular arthritis†, n (%)	14 (12.3)	13 (23.6)	7 (25.9)	5 (33.3)	98 (81.7)
HAQ, mean (SD)	0.72 (0.6)	0.72 (0.6)	0.62 (0.5)	0.65 (0.5)	1.07 (0.7)
DAS28, mean (SD)	3.6 (1.2)	4.0 (1.2)	4.1 (1.1)	4.1 (1.1)	5.4 (1.3)
Criteria points, median (25, 75 percentile)	3 (1, 4)	4 (2, 4)	4 (3, 4)	4 (3, 4)	7 (6, 9)

Small joint involvement, metacarpophalangeal joints, proximal interphalangeal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints and wrists.
*Fulfilling the erosion criterion.
†>4 swollen joints.
28-TJC, 28-tender joint count; 68-SJC, 68-swollen joint count (standard 66-SJC plus hip joints); ACPA, anticitrullinated peptide antibody; CRP, C reactive protein; DAS28, Disease Activity Score-28; ESR, erythrocyte sedimentation rate; HAQ, Health Assessment Questionnaire; RA, rheumatoid arthritis; RF, rheumatoid factor.

weeks. A proportion of patients have erosions at first presentation,[6] and this has been found to be a predictor for severe destructive disease in patients with early RA.[7 8] However, data from our study regarding DMARD use and final clinical diagnosis indicate that the majority of patients fulfilling only the erosion criterion at baseline were false-positives.

To our knowledge, few studies have looked at the role of the erosion criterion in classifying patients with early RA. Le Loët *et al*[9] studied 310 patients with early arthritis (median symptom duration of 128 days), and no patients with a 2010 RA criteria score <6 fulfilled the erosion criterion at baseline. Patients in our study had a higher mean Disease Activity Score-28 at baseline than the French patients, which could have contributed to the discrepancy in the erosive findings. The patients in the present study who solely fulfilled the erosion criterion at baseline were all seronegative, had fewer involved joints and more often male compared with patients fulfilling the 2010 RA criteria. Most of these are expected, as being seropositive and having more involved joints are included as points in the 2010 RA criteria.

Table 2 Distribution of erosive joints in patients not fulfilling the 2010 RA criteria at baseline

	Erosive joints at baseline			
	PIP	MCP	Wrist*	MTP†
≥1 Erosive joint (n=55)	23	17	20	34
≥2 Erosive joints (n=27)	11	13	18	20
≥3 Erosive joints (n=15)	6	8	15	15

*The CMC bone, the trapezium, the scaphoid, the lunate, the radial and the ulnar bone.
†IP1, MTP2, MTP3, MTP4 and MTP5.
RA, rheumatoid arthritis; PIP, proximal interphalangeal; MCP, metacarpophalangeal; CMC, carpometacarpal; IP1, 1st interphalangeal.

The MTP joints in the feet had the highest occurrence of erosions both at baseline and during follow-up. A previous study of patients with UA with erosive joints at baseline suggested that presence of erosions in the joints of the feet was slightly more predictive for developing RA than erosions in the hand joints.[10]

The task force that defined the erosion criterion concluded that the specificity of a cut-off of ≥2 erosive joints would be too low.[3] Our results show that using ≥2 erosive joints as cut-off would have increased the number of patients being classified as RA with 27 (16.0%) at baseline and another 4 patients during follow-up. Although the baseline characteristics of patients with ≥2 and ≥3 were quite similar, the low number of patients with radiographic follow-up data makes it difficult to consider the consequences of having ≥2 erosive joints as cut-off regarding disease course and outcome.

A limitation of our study is the low rate of radiographic follow-up. We believe that many patients declined or were not referred to radiographic follow-up because they were feeling healthy. Another limitation is the small number of patients with erosive joints in general, which precludes meaningful statistical comparisons. Additionally, if our inclusion criteria had allowed for longer duration of joint swelling, the proportion of patients developing ≥3 erosive joints might have been larger, as RA often has an insidious onset. Ideally, the baseline radiographs should also have been read separately; however, we do not think this has had a major impact on the results because the reader was unaware of the purpose of scoring the radiographs. Additionally, there were only a few patients developing incident erosions.

In conclusion, few patients were classified as RA based on the erosion criterion without fulfilling the 2010 RA criteria. Of those who did, almost all did so at baseline; thus, our results suggest that follow-up radiographs in patients with early UA might be of limited value for classifying patients with RA. Furthermore,

data regarding DMARD use and clinical diagnosis indicate that despite having erosions at baseline, patients with UA may end up with other clinical diagnoses than RA.

Acknowledgements The authors thank the patients, the physicians and study nurses, and the laboratory and radiology staff for participating in the NOR-VEAC study.

Contributors GHB, ESN, MDM, DvdH, TKK and EL contributed to the study design, including formulation of the research questions. GHB, ESN, MDM and EL were responsible for analysing and interpreting the data, and drafting the manuscript. All authors critically revised the manuscript and approved the final version.

Competing interests None declared.

Patient consent Obtained.

Ethics approval The Regional Ethics Committee of Southern Norway.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 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.
2 Knevel R, Lukas C, van der Heijde D, *et al*. Defining erosive disease typical of RA in the light of the ACR/EULAR 2010 criteria for rheumatoid arthritis; results of the data driven phase. *Ann Rheum Dis* 2013;72:590–5.
3 van der Heijde D, van der Helm-van Mil AH, Aletaha D, *et al*. EULAR definition of erosive disease in light of the 2010 ACR/EULAR rheumatoid arthritis classification criteria. *Ann Rheum Dis* 2013;72:479–81.
4 Mjaavatten MD, Haugen AJ, Helgetveit K, *et al*. Pattern of joint involvement and other disease characteristics in 634 patients with arthritis of less than 16 weeks' duration. *J Rheumatol* 2009;36:1401–6.
5 van der Heijde D. How to read radiographs according to the sharp/van der Heijde method. *J Rheumatol* 2000;27:261–3.
6 Aletaha D, Funovits J, Breedveld FC, *et al*. Rheumatoid arthritis joint progression in sustained remission is determined by disease activity levels preceding the period of radiographic assessment. *Arthritis Rheum* 2009;60:1242–9.
7 Smolen JS, van der Heijde DM, Aletaha D, *et al*. Progression of radiographic joint damage in rheumatoid arthritis: independence of erosions and joint space narrowing. *Ann Rheum Dis* 2009;68:1535–40.
8 van Nies JA, van Steenbergen HW, Krabben A, *et al*. Evaluating processes underlying the predictive value of baseline erosions for future radiological damage in early rheumatoid arthritis. *Ann Rheum Dis* 2015;74:883–9.
9 Le Loët X, Nicolau J, Boumier P, *et al*. Validation of the 2010-ACR/EULAR -classification criteria using newly EULAR-defined erosion for rheumatoid arthritis on the very early arthritis community-based (VErA) cohort. *Joint Bone Spine* 2015;82:38–41.
10 Thabet MM, Huizinga TW, van der Heijde DM, *et al*. The prognostic value of baseline erosions in undifferentiated arthritis. *Arthritis Res Ther* 2009;11:R155.

EXTENDED REPORT

Molecular basis for increased susceptibility of Indigenous North Americans to seropositive rheumatoid arthritis

Stephen W Scally,[1] Soi-Cheng Law,[2] Yi Tian Ting,[1] Jurgen van Heemst,[3] Jeremy Sokolove,[4] Aaron J Deutsch,[5] E Bridie Clemens,[6] Antonis K Moustakas,[7] George K Papadopoulos,[8] Diane van der Woude,[3] Irene Smolik,[9] Carol A Hitchon,[9] David B Robinson,[9] Elizabeth D Ferucci,[10] Charles N Bernstein,[9] Xiaobo Meng,[9] Vidyanand Anaparti,[9] Tom Huizinga,[3] Katherine Kedzierska,[6] Hugh H Reid,[1] Soumya Raychaudhuri,[11,12,13,14] René E Toes,[3] Jamie Rossjohn,[1,15,16] Hani El-Gabalawy,[9] Ranjeny Thomas[2]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-211300).

For numbered affiliations see end of article.

Correspondence to
Professor Ranjeny Thomas, University of Queensland Diamantina Institute, Level 4, R Wing, Princess Alexandra Hospital, Ipswich Road, Woolloongabba 4102, Queensland, Australia; ranjeny.thomas@uq.edu.au

SWS and S-CL contributed equally.

Received 13 February 2017
Revised 26 May 2017
Accepted 1 July 2017
Published Online First 11 August 2017

ABSTRACT

Objective The pathogenetic mechanisms by which *HLA-DRB1* alleles are associated with anticitrullinated peptide antibody (ACPA)-positive rheumatoid arthritis (RA) are incompletely understood. RA high-risk *HLA-DRB1* alleles are known to share a common motif, the 'shared susceptibility epitope (SE)'. Here, the electropositive P4 pocket of HLA-DRB1 accommodates self-peptide residues containing citrulline but not arginine. HLA-DRB1 His/Phe13β stratifies with ACPA-positive RA, while His13βSer polymorphisms stratify with ACPA-negative RA and RA protection. Indigenous North American (INA) populations have high risk of early-onset ACPA-positive RA, whereby HLA-DRB1*04:04 and HLA-DRB1*14:02 are implicated as risk factors for RA in INA. However, HLA-DRB1*14:02 has a His13βSer polymorphism. Therefore, we aimed to verify this association and determine its molecular mechanism.

Methods HLA genotype was compared in 344 INA patients with RA and 352 controls. Structures of HLA-DRB1*1402-class II loaded with vimentin-64Arg$_{59\text{-}71}$, vimentin-64Cit$_{59\text{-}71}$ and fibrinogen β–74Cit$_{69\text{-}81}$ were solved using X-ray crystallography. Vimentin-64Cit$_{59\text{-}71}$-specific and vimentin$_{59\text{-}71}$-specific CD4+ T cells were characterised by flow cytometry using peptide-histocompatibility leukocyte antigen (pHLA) tetramers. After sorting of antigen-specific T cells, TCRα and β-chains were analysed using multiplex, nested PCR and sequencing.

Results ACPA$^+$ RA in INA was independently associated with *HLA-DRB1*14:02*. Consequent to the His13βSer polymorphism and altered P4 pocket of HLA-DRB1*14:02, both citrulline and arginine were accommodated in opposite orientations. Oligoclonal autoreactive CD4+ effector T cells reactive with both citrulline and arginine forms of vimentin$_{59\text{-}71}$ were observed in patients with HLA-DRB1*14:02$^+$ RA and at-risk ACPA$^-$ first-degree relatives. HLA-DRB1*14:02-vimentin$_{59\text{-}71}$-specific and HLA-DRB1*14:02-vimentin-64Cit$_{59\text{-}71}$-specific CD4+ memory T cells were phenotypically distinct populations.

Conclusion HLA-DRB1*14:02 broadens the capacity for citrullinated and native self-peptide presentation and T cell expansion, increasing risk of ACPA+ RA.

To cite: Scally SW, Law S-C, Ting YT, *et al*. *Ann Rheum Dis* 2017;**76**:1915–1923.

INTRODUCTION

Rheumatoid arthritis (RA) is an autoimmune disease with peak incidence in the sixth decade and prevalence of 1% in Caucasians, linked to *HLA-DRB1*. *HLA-DRB1* alleles associated with anticitrullinated peptide antibody (ACPA)-positive RA in Caucasians share a common motif in the third hypervariable region, the 'shared susceptibility epitope (SE)', which was shown to accommodate citrullinated (Cit) self-epitopes.[1–3] Citrulline is post-translationally modified from arginine during inflammation, endoplasmic reticulum (ER) stress and autophagy.[4 5] Various RA-associated Cit-autoantigens, recognised by ACPA, are present in inflamed sites, including joint tissues. Low-titre ACPA develops in healthy individuals associated with environmental risk factors including smoking and periodontitis, but are unrelated to *HLA-DR SE*.[6–11] In the immediate pre-RA period, ACPA isotype diversity and titre increase — a process associated with antigen-specific CD4+ T cell help for affinity maturation in germinal centres.[8] *HLA-DR SE* is associated with ACPA$^+$ RA rather than ACPA, implying that presentation of Cit-autoantigens bound to HLA-DR SE molecules to CD4$^+$ T cells is associated with RA development in at-risk individuals carrying *HLA-DR SE*.

Based on genome-wide studies (GWAS) in patients of predominantly Caucasian and Asian ethnicity, *HLA* alleles associated with ACPA-positive RA, including *HLA-DRB1*04:01*, *HLA-DRB1*04:05* and *HLA-DRB1*01:01* (ORs of 2.17–4.44), were found to share a common motif at amino acid positions 11, 13, 71 and 74, influencing the P4 antigen-binding pocket of DRβ.[2] Moreover, Val11βSer and His13βSer polymorphisms within that motif were found in genomic studies to stratify with ACPA-negative RA in Caucasians.[3] The discovery that P4-Cit was accommodated but the positively charged Arg was excluded from the electropositive P4 pocket of HLA-DRB1*04:01/04 suggested that preferential presentation of Cit-autoantigens might underpin the association of ACPA$^+$ RA with the SE.[12] While HLA-DRB1*04:01-restricted CD4$^+$ memory T cells recognising Cit-autoantigens have been reported in patients with RA,[12 13] their role

in disease development is unclear. To date it has not been determined how autoreactive T cells respond to Cit-autoantigen in RA and whether they undergo clonal expansion due to antigen experience in patients with RA or in HLA-SE$^+$ at-risk first-degree relatives (FDRs).

Furthermore, in unravelling antigen presentation by HLA-DR molecules in RA, the ethnic mix of the samples included in GWAS may skew interpretations of amino acids contributing to binding motifs made from genomic studies. The Indigenous North American (INA) population has a twofold to threefold higher prevalence of RA than Caucasians, with RA onset peaking earlier in the fourth decade of life.[14] Moreover, 90% of INA patients with RA are ACPA$^+$. In this population, FDRs have a high prevalence of joint symptoms and of ACPA positivity, predisposing them to RA.[15 16] To date, the *HLA* association with RA in INA has been sought in studies of <100 patients.[17–19] These studies suggested that the *HLA-SE* alleles also predispose to RA in INA. *HLA-DRB1*04:04* is the most frequent *SE* allele in these populations, followed by *HLA-DRB1*14:02*. Since HLA-DRB1*14:02 carries the β13Ser residue, which was interpreted to stratify with ACPA-negative RA,[3] we investigated the genetic and underlying molecular bases for the increased risk of severe ACPA$^+$ RA in INA.

MATERIALS AND METHODS

Study participants

RA cases, non-RA controls and FDRs were recruited from INA populations in Central Canada (Cree, Ojibway and Ojicree) and Alaska Native people (from Southcentral and Southeast Alaska). DNA for HLA typing, serum and peripheral blood mononuclear cells (PBMC) were isolated.

HLA-DRB1*14:02 expression and purification

Peptide-loaded HLA-DRB1*14:02 molecules were purified, crystallised and structures determined.

Multiplex ACPA assay

Serum levels of antibodies targeting 40 putative RA-associated autoantigens were measured using a custom bead-based immunoassay on a Bio-Plex platform, as previously described.[20]

Tetramer staining and analysis of TCR repertoire

Tetramer staining and T cell receptor (TCR) repertoire analysis used previously published methods, with some modifications.[12 21 22]

Details of all methods are available in online supplementary methods, supplementary table 1 and supplementary figure 1.

RESULTS

*HLA-DRB1*14:02* is independently associated with ACPA$^+$ RA in INA

Although very rare in Caucasians and Asians, *HLA-DRB1*14:02* has a prevalence of up to 80% in some INA populations, suggesting a particular survival advantage against pathogens.[17–19] To determine the *HLA-DRB1* association with ACPA$^+$ RA in INA, we genotyped the largest cohort available, comprising 344 INA patients with RA and 352 controls. Rheumatoid factor (RF) and ACPA status was known in 241/344 patients with RA: 90% were seropositive (RF and/or ACPA$^+$). In patients with seropositive RA, 32% carried *HLA-DRB1*14:02* and 45% *HLA-DRB1*04:04*. One-third of patients carrying *HLA-DRB1*14:02* carried an additional *SE* allele. In healthy controls (HCs), 28% carried *DRB1*14:02*, 22% *HLA-DRB1*04:04* and 17% carried an additional *SE* allele. After stratifying patients with RA according to HLA status, *HLA-DRB1*14:02* was a risk factor for seropositive RA (OR 2.38) independent of other *SE* alleles (table 1). In INA patients with RA, the most commonly associated other *SE* allele is *HLA-DRB1*04:04* (table 1). In INA, RA risk was associated with *HLA-DRB1* alleles with a conserved SE motif at 71 and 74 in all RA (OR=2.48, 95% CI 1.70 to 3.60, p<0.0001) and in seropositive patients with RA (OR=2.46, 95% CI 1.58 to 3.81, p=0.0001) (table 1). We note that without genome-wide genotyping, we cannot rule out the possibility of confounding due to case–control differences in ancestry.

To stratify ACPA response with genotype, sera of 232 INA patients with RA were tested in a multiplex ACPA antigen array and given an ACPA score (sum of all normalised ACPA titres divided by number of epitopes).[23] ACPA score was higher in INA patients with RA who were either *HLA-DRB1*14:02* homozygotes or *HLA-DRB1*14:02/HLA-SE* compound heterozygotes than those who were *HLA-SE*-negative (p<0.05). ACPA scores in *HLA-DRB1*14:02*$^+$ patients were equivalent to those in *HLA-SE*$^+$ patients lacking *HLA-DRB1*14:02*. ACPA specificities increased among *HLA-DRB1*14:02*$^+$ patients and included vimentin-64Cit$_{58\text{-}77}$, filaggrin-56Cit$_{48\text{-}65}$ and fibrinogen-α−573Cit$_{556\text{-}575}$ (figure 1A,B). Thus, despite polymorphisms of Val11βSer and His13βSer, INA individuals carrying *HLA-DRB1*14:02* develop a broad ACPA response whether or not they carry other *SE* alleles. This implies binding and presentation of a variety of Cit-autoantigens by *HLA-DRB1*14:02* to autoreactive T cells.

Accommodation of arginine and citrulline residues within the P4 pocket of HLA-DRB1*14:02

The HLA-DRB1 chain contains 12 polymorphic residues that have been directly implicated in peptide binding.[24] HLA-DRB1*14:02 differs from HLA-DRB1*04:01, HLA-DRB1*04:04 or HLA-DRB1*01:01 in eight of these residues, which shape P4, P6, P7 and P9 pockets (figure 2A). To test whether HLA-DRB1*14:02 presents autoantigens differently from HLA-DRB1*04:01, HLA-DRB1*04:04 or HLA-DRB1*01:01, we compared the capacity of each HLA binding pocket to accommodate Cit and Arg residues using the influenza-derived HA$_{305\text{-}319}$ peptide.[25–28] Conversion of Arg to Cit at peptide positions interacting with P4 enhanced

Table 1 Association of HLA-DRB1*14:02 with all RA and seropositive RA in INA patients with RA and controls

	HC, n (%)	All RA, n (%)	OR (CI)	p	Seropositive RA, n (%)	OR (CI)	p
SE$^-$	106 (30)	51 (15)	Ref		32 (15)	Ref	
14:02$^+$/Other SE$^-$	64 (18)	80 (23)	2.60 (1.63 to 4.15)	0.0001	46 (21)	2.38 (1.38 to 4.12)	0.0019
14:02$^-$/Other SE$^+$	147 (42)	178 (52)	2.52 (1.69 to 3.75)	<0.0001	117 (54)	2.64 (1.66 to 4.19)	<0.0001
14:02$^+$/Other SE$^+$	35 (10)	35 (10)	2.08 (1.17 to 3.70)	0.0127	23 (11)	2.18 (1.13 to 4.20)	0.02
Any SE$^+$	246 (70)	293 (85)	2.48 (1.70 to 3.60)	<0.0001	186 (85)	2.46 (1.58 to 3.81)	0.0001

Patients and controls were stratified according to the presence of SE. SE-positive individuals were further stratified as shown. Other SE alleles included DRB1*01:01, DRB1*01:02, DRB1*04:01, DRB1*04:04, DRB1*04:05, DRB1*04:08, DRB1*04:10, DRB1*04:13 and DRB1*10:01.
HC, healthy controls; INA, Indigenous North American; RA, rheumatoid arthritis; SE, susceptibility epitope.

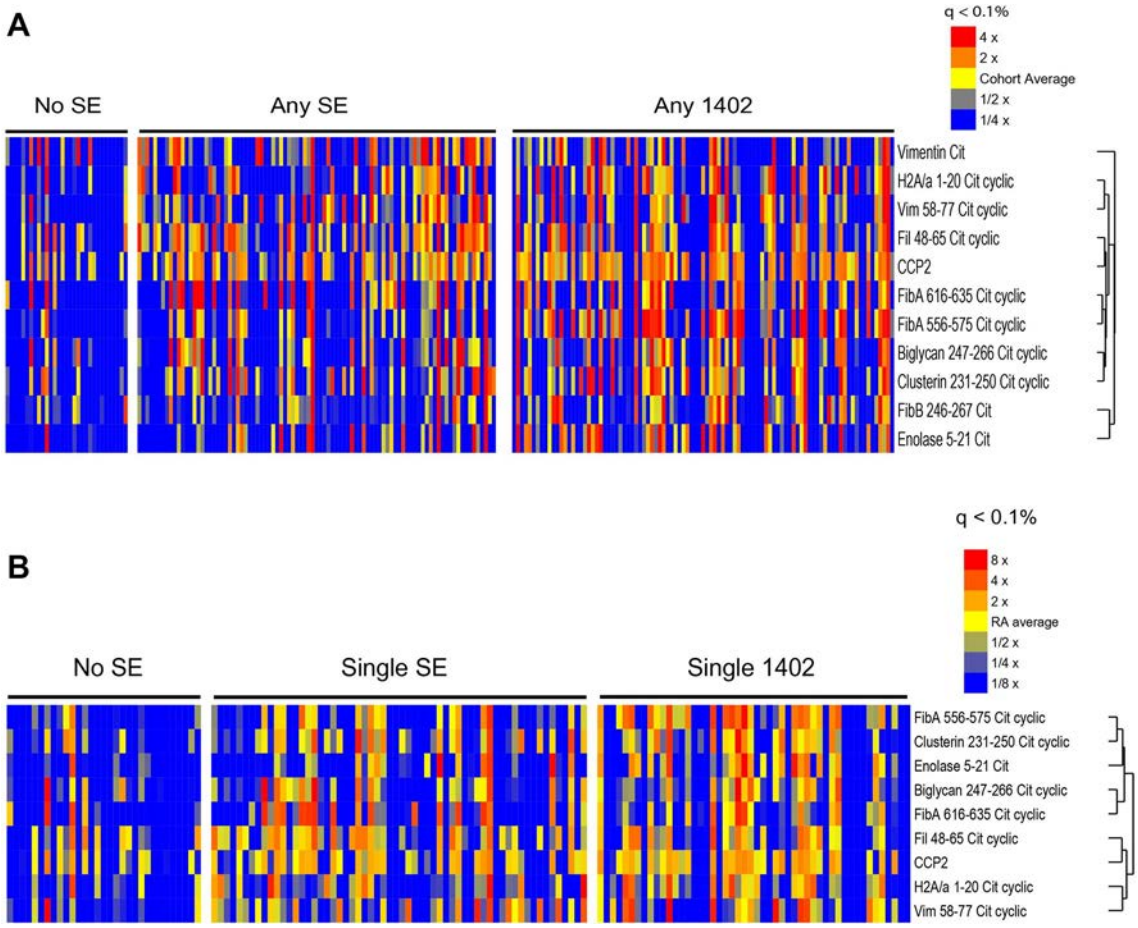


Figure 1 HLA-DRB1*14:02 is associated with a broad ACPA response. Serum levels of antibodies targeting RA-associated Cit-autoantigens were measured in serum from 232 INA patients with RA using a custom bead-based fluorescence immunoassay. Fluorescence intensities for 10 Cit-autoantigens and CCP were clustered (A) according to the presence or absence of 1 or 2 HLA-DR-SE alleles or HLA-DRB1*14:02 with any HLA-DR-SE allele, and (B) according to presence or absence of a single HLA-DR-SE or HLA-DRB1*14:02 allele, as shown. Each column represents one patient sample. ACPA, anticitrullinated peptide antibody; CCP, cyclic citrullinated peptide; Cit, citrullinated; RA, rheumatoid arthritis; SE, susceptibility epitope.

peptide binding affinity to HLA-DRB1*01:01 by twofold and to HLA-DRB1*04:01 by tenfold. Conversion of Arg to Cit at peptide positions interacting with the P6 and P9 pockets enhanced peptide binding affinity to HLA-DRB1*01:01, and P7 to HLA-DRB1*04:01. In contrast, peptides containing an Arg or Cit residue at positions interacting with P4 of HLA-DRB1*14:02 had similar binding affinity, and peptides containing Arg at positions interacting with P6 and P9 had increased affinity relative to Cit peptides (figure 2B). While the IC50 of the self-peptide, vimentin-64Cit$_{59\text{-}71}$, was decreased by 1.3-fold relative to the 64-Arg variant for HLA-DRB1*01:01, and the 64-Arg variant did not bind HLA-DRB1*04:01, HLA-DRB1*14:02 bound both 64-Arg and 64-Cit variants with IC50 of 19 µM (figure 2C), indicating that both residues could be accommodated within P4 of HLA-DRB1*14:02, or that HLA-DRB1*14:02 presented in differing peptide binding registers.

Structural basis of peptide presentation by HLA-DRB1*14:02

We solved the structure of HLA-DRB1*14:02 in complex with vimentin-64Arg$_{59\text{-}71}$, vimentin-64Cit$_{59\text{-}71}$ and fibrinogen β74Cit$_{69\text{-}81}$ (online supplementary table 3, figure 3A–D). These HLA-DRB1*14:02 structures overlaid closely (figure 3E), ruling out markedly differing binding modes to accommodate these differing epitopes. Within the HLA-DRB1*14:02-vimentin-64Cit$_{59\text{-}71}$ and HLA-DRB1*14:02-vimentin-64Arg$_{59\text{-}71}$ structures, Tyr, Ser and Arg occupied the P1, P6 and P9 pockets of HLA-DRB1*14:02, respectively (figure 3B,C). In the HLA-DRB1*14:02-fibrinogen β74Cit$_{69\text{-}81}$ complex, Tyr, Ala and Ala occupied the P1, P6 and P9 pockets of HLA-DRB1*14:02, respectively (figure 3D). The largest structural differences between the peptides in each binary complex centred on the residue occupying the P4 pocket of HLA-DRB1*14:02, namely P4-Cit and P4-Arg in vimentin-64Cit$_{59\text{-}71}$, fibrinogen β74Cit$_{69\text{-}81}$ and vimentin-64Arg$_{59\text{-}71}$, respectively. These structures clearly show that P4-Arg and P4-Cit are presented in alternative orientations, whereby the P4-Arg projects inwards, whereas the P4-Cit projects outwards from the HLA-DRB1*14:02 Ag-binding cleft (figure 3F–H). The P4-Arg is buried in the pocket to avoid interactions with the positively charged β71Arg residue. This orientation is promoted by β11Ser and β13Ser, whereupon these two small polar residues allow the accommodation of Arg by providing the necessary space and H-bonding partners (figure 3F). Larger residues in HLA-DRB1*04:01 (β11Val and β13His) and HLA-DRB1*01:01 (β11Leu and β13Phe) as well as the charge repulsion of β13His in HLA-DRB1*04:01 would prevent the accommodation of Arg at P4.[12] The P4-Arg residue is stabilised by H-bonds with Ser11β, Ser13β and Tyr30β, and a salt bridge with Glu28β (figure 3F). In contrast, the P4-Cit sits upright in both Cit-epitopes, similar to its orientation in P4-Cit from Cit epitopes presented by HLA-DRB1*04:01/04:04

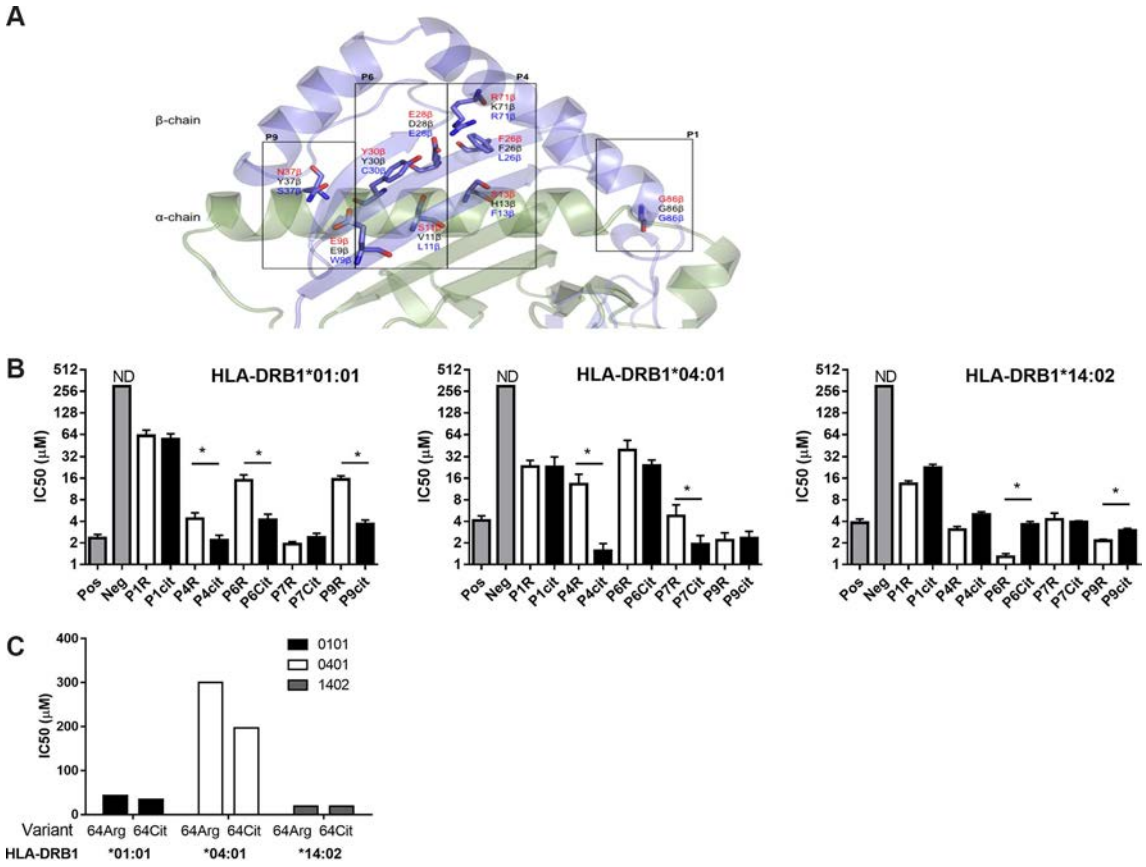


Figure 2 Accommodation of arginine and citrulline residues by HLA-DRB1*01:01, HLA-DRB1*04:01 and HLA-DRB1*14:02. (A) The peptide binding groove of an HLA-DR molecule is shown in cartoon with polymorphic residues shown as sticks corresponding to the residues present in HLA-DRB1*14:02, which is associated with RA among the INA. Schematic representation of the differences in pockets between HLA-DRB1*14:02 (red font), HLA-DRB1*04:01 (black font) and HLA-DRB1*01:01 (blue font). (B) Competitive binding of a biotin-labelled $HA_{305-319}$ peptide with an unlabelled $HA_{305-319}$ peptide or $HA_{305-319}$ variants with citrulline or arginine residues in P1, P4, P6, P7 and P9 to HLA-DRB1*01:01, HLA-DRB1*04:01 and HLA-DRB1*14:02. ND, non-detectable binding affinity. Pooled binding data from at least three experiments are shown and error bars depict the variation between experiments. (C) Competitive binding of a biotin-labelled $HA_{305-319}$ peptide with unlabelled $vimentin_{59-71}$ and $vimentin\text{-}64Cit_{59-71}$ to HLA-DRB1*01:01, HLA-DRB1*04:01 and HLA-DRB1*14:02. INA, Indigenous North American; RA, rheumatoid arthritis.

(figure 3G,H).[12] The P4-Cit is stabilised by H-bonds with Arg71β and Gln70β (figure 3G,H). Accordingly, we demonstrate a conserved positioning of the citrulline residue in two distinct epitopes that contrast the orientation of the non-Cit residue within the P4 pocket. Similar to its orientation in HLA-DRB1*04:01 and HLA-DRB1*04:04,[12 29] P4-Cit was solvent-exposed in HLA-DRB1*14:02 and could potentially interact with TCR.

Antigen-experienced, oligoclonally expanded T cells recognise Arg and Cit variants of $vimentin_{59-71}$ presented by HLA-DRB1*14:02

Although both vimentin-64-Arg and 64-Cit variant peptides bound HLA-DRB1*14:02, only P4-Cit sits upright, and thereby potentially able to interact with the TCR. Thus, we addressed whether autoreactive T cells with TCRs recognising one or both epitopes were present in the periphery and displayed evidence of in vivo expansion in response to antigen presentation. We analysed T cells recognising HLA-DRB1*14:02-vimentin-$64Cit_{59-71}$ and HLA-DRB1*1402-$vimentin_{59-71}$ tetramers in 10 HLA-DRB1*14:02$^+$ INA patients with RA, 10 HLA-DRB1*14:02$^+$ ACPA$^-$ FDRs and 6 HLA-DRB1*14:02$^+$ non-INA HC subjects. FDRs in the INA population have a high burden of environmental risk factors for RA, a high level of background HLA-DR SE genes, high inflammatory C-reactive protein (CRP) and a high prevalence of joint symptoms.[16] Among the 13 FDRs studied (online supplementary table 3), 85% were past smokers and 31% had an abnormal CRP >8. Therefore we could compare T cells recognising vimentin or Cit-vimentin in individuals both with RA and with high risk of future RA. Similar frequencies of CD4$^+$ T cells recognised vimentin 64-Arg and 64-Cit variant peptides in HLA-DRB1*14:02$^+$ FDRs and patients with RA, and with similar tetramer staining intensity (mean fluorescence intensity (MFI)) (figure 4A,B). In all individuals, HLA-DRB1*14:02-vimentin-$64Cit_{59-71}$-reactive and HLA-DRB1*14:02-$vimentin_{59-71}$-reactive CD4$^+$ T cells were significantly enriched in CD25$^+$CD127$^+$ effector (Teff) and CD25$^+$CD127$^-$regulatory T cells (Treg)[30] compared with the total CD4$^+$ PB T cell pool (Treg $p<0.0001$, Teff $p<0.05$; figure 4C), consistent with antigen experience in vivo. This enrichment did not differ between patients with RA and FDRs, indicating that antigen experience and formation of memory develop before the onset of ACPA in INA. The numbers of circulating HLA-DRB1*14:02-vimentin-$64Cit_{59-71}$-reactive and HLA-DRB1*14:02-$vimentin_{59-71}$-reactive CD4$^+$ T cells were correlated in each individual patient with RA and each FDR ($r^2=0.74$, $p<0.0001$; figure 4D).

T cell effector function is balanced by the suppressive activity of Treg.[31] The ratio of

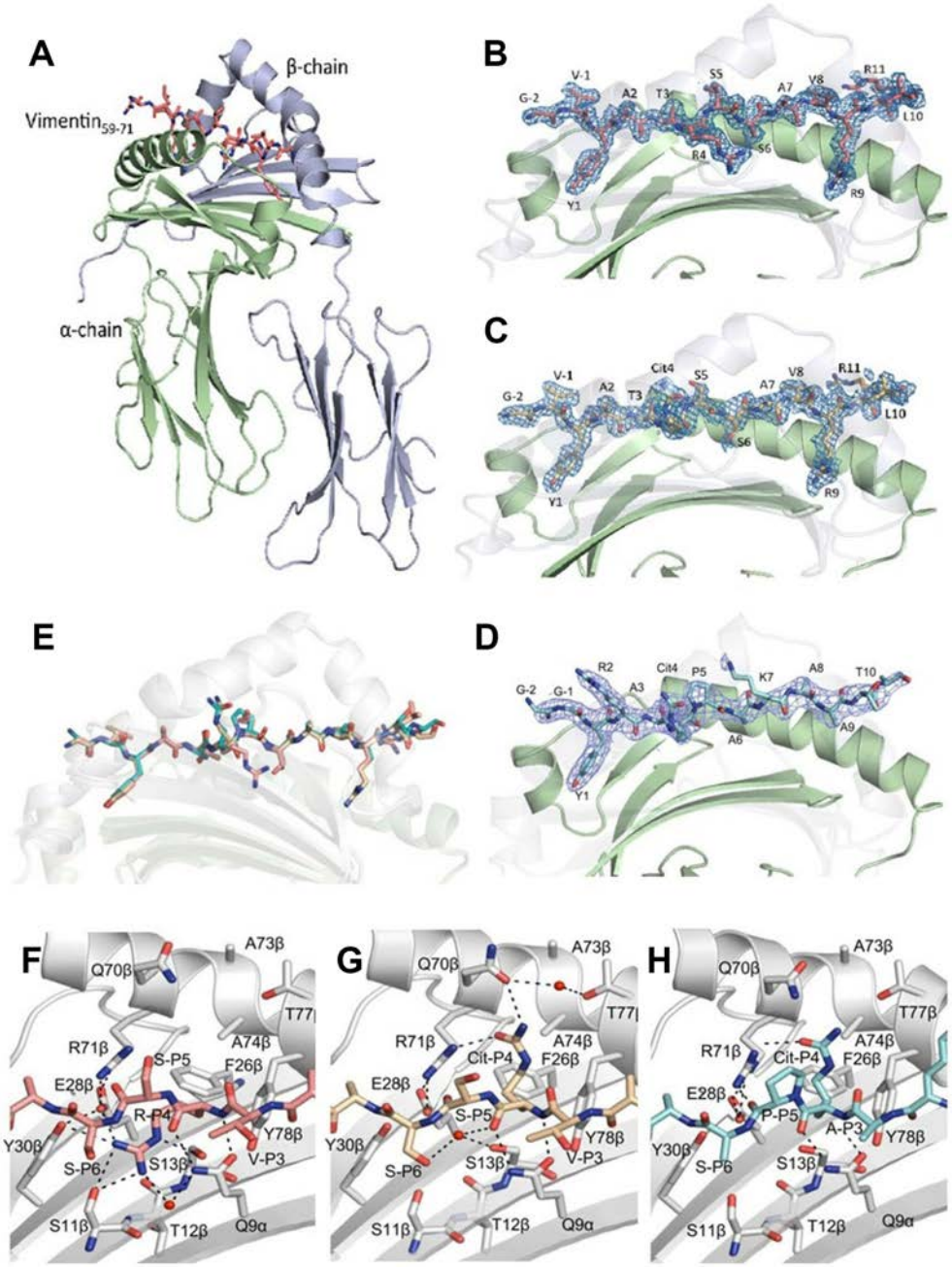


Figure 3 Crystal structure of HLA-DRB1*1402 in complex with vimentin$_{59\text{-}71}$, vimentin-64Cit$_{59\text{-}71}$ and fibrinogen β 74Cit$_{69\text{-}81}$.The α-chains and β-chains of (A) the HLA-DRB1*1402 in complex with vimentin$_{59\text{-}71}$ are shown in cartoon representation and coloured in light green and light blue, respectively, while peptide is displayed as sticks. Side view of (B) HLA-DRB1*1402-vimentin$_{59\text{-}71}$, (C) HLA-DRB1*1402-vimentin-64Cit$_{59\text{-}71}$ and (D) HLA-DRB1*1402-fibrinogen β74Cit$_{69\text{-}81}$. The carbons are coloured deep salmon, light orange and teal for vimentin$_{59\text{-}71}$, vimentin-64Cit$_{59\text{-}71}$ and fibrinogen β74Cit$_{69\text{-}81}$, respectively; nitrogens are coloured in blue and oxygens are coloured in red. The β-chain is transparent to help visualise the peptide. The unbiased 2Fo-Fc electron density map of the peptides is shown in blue and contoured to 1 σ. (E) Superposition of the three peptides bound to HLA-DRB1*1402. The P4 pocket of HLA-DRB1*1402 bound to (F) vimentin$_{59\text{-}71}$, (G) vimentin-64Cit$_{59\text{-}71}$ and (H) fibrinogenβ74Cit$_{69\text{-}81}$. The P4-Arg in the vimentin$_{59\text{-}71}$ peptide is buried in the P4 pocket. The P4-Cit of the vimentin-64Cit$_{59\text{-}71}$ and fibrinogenβ74Cit$_{69\text{-}81}$ peptide adopts an upright conformation.

HLA-DRB1*14:02-vimentin$_{59\text{-}71}$Cit-64-reactive and HLA-DRB1*14:02-vimentin$_{59\text{-}71}$-reactive Teff/Treg was significantly lower than that of the total CD4$^+$ T cell pool in patients with RA and FDRs ($p<0.001$, figure 4E), consistent with active regulation of the autoreactive T cells. To understand the particular role of CD4$^+$ T cells of each vimentin specificity further, we analysed an additional six HLA-DRB1*14:02$^+$ INA individuals (three with RA, three FDRs; clinical details in online supplementary table 3) and six HLA-DRB1*14:02$^+$ HC for markers of memory T cell activation and differentiation relative to total CD4+ T cells. In INA patients with RA and FDRs, the proportion of CD28$^+$ memory T cells was significantly higher among vimentin$_{59\text{-}71}$-reactive CD4$^+$ T cells ($p<0.05$; figure 4F). Moreover, cells expressing CD69 were significantly enriched among vimentin$_{59\text{-}71}$-reactive CD4$^+$ T cells in these individuals. Some CD69$^+$ vimentin$_{59\text{-}71}$-reactive

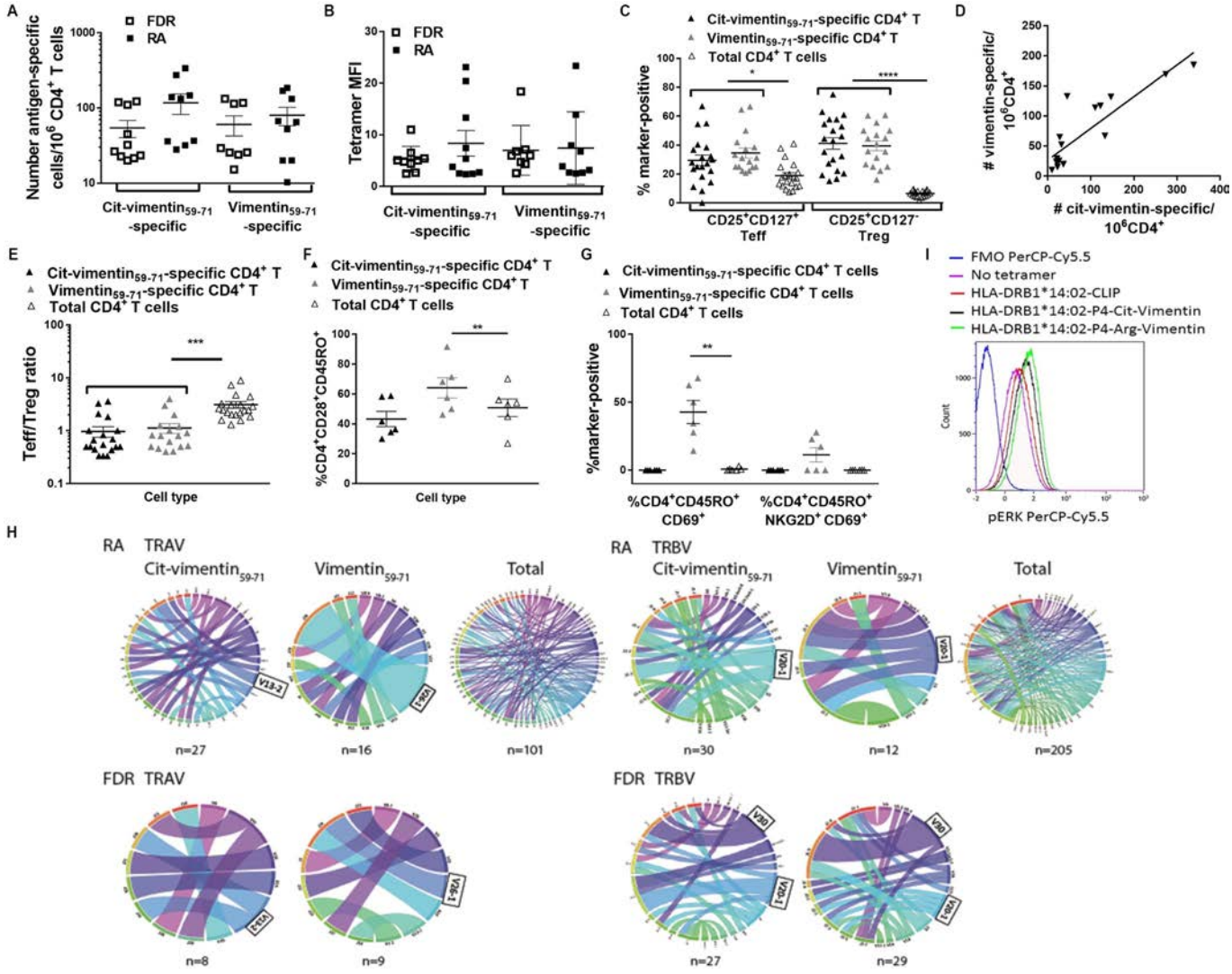


Figure 4 Presentation of vimentin$_{59-71}$ and vimentin-64Cit$_{59-71}$ self-antigens in context of HLA-DRB1*14:02 to CD4$^+$ T cells in individuals with and at high risk of ACPA$^+$ RA. (A–D) PBMC from HLA-DRB1*14:02$^+$ RA patients W1:1-10 and HLA-DRB1*14:02$^+$ FDR W1:1-10 were stained with PE-labelled HLA-DRB1*1402-vimentin64Cit$_{59-71}$ or HLADRB1*14:02-vimentin$_{59-71}$ tetramer, CD4-APC, CD25-PE/Cy7, and CD127-BV421. Live CD4+ FITC-lineage-negative tetramer+ T cells were gated based on an FMO sample stained with unlabelled primary and conjugated secondary antibodies. The number calculated relative to the CD4+ T cell count, (A) and mean fluorescence intensity (MFI) (B) of tetramer$^+$ cells within the total CD4$^+$ population is shown for RA patients and FDR; (C) the proportion of CD4$^+$CD25$^+$CD127$^+$ Teff and CD4$^+$CD25$^+$CD127$^-$ Treg within vimentin64Cit$_{59-71}$-reactive and vimentin$_{59-71}$-reactive CD4+ T cell gates and the total CD4$^+$ gate is shown for pooled RA patients and FDR. (D) The numbers of vimentin64Cit$_{59-71}$-reactive and vimentin$_{59-71}$-reactive CD4$^+$ T cells were correlated in pooled RA patients and FDR (R^2 0.74, p<0.0001). (E) The Teff/Treg ratio of vimentin64Cit$_{59-71}$ and vimentin$_{59-71}$-reactive CD4$^+$ T cells is plotted relative to total CD4$^+$ T cell ratio in pooled RA patients and FDR; (F) CD45RO$^+$CD28$^+$, (G) CD45RO$^+$CD69$^+$ and CD45RO$^+$CD69$^+$ NKG2D$^+$ vimentin$_{59-71}$- reactive CD4$^+$, vimentin64Cit$_{59-71}$-reactive CD4$^+$ and total CD4$^+$ T cells are plotted in pooled RA patients W2:1-3 and FDR W2:1-3. (H) Frequencies of V-J pairing for TCRα and TCRβ among vimentin64Cit$_{59-71}$-specific and vimentin$_{59-71}$-specific CD4$^+$ T cells from HLADRB1*14:02$^+$ RA patients (n=5) and FDR (n=5), and tetramer-negative CD3$^+$ T cells sorted from DRB1*14:02$^+$ RA patients (n=3). The width of the bands is proportional to the frequency of TCR sequences with a particular V-J pairing. Details of productive TCRα and TCRβ sequences in online supplementary table 2 and 4. The figures were generated using the Circos online Table Viewer software (http://mkweb.bcgsc.ca/tableviewer/). (I) 10^6 P2F3 SKW3 cells were stimulated with 5 ng/mL unlabelled HLA-DRB1*14:02-vimentin$_{59-71}$ or HLA-DRB1*14:02-vimentin-64Cit$_{59-71}$ tetramer for 4 hours at 37°C. Cells were fixed and permeabilised, then stained with PerCP/ Cy5.5-ERK1/2 and PB-CD3. Live CD3+ cells expressing the GFP marker gene were gated. Representative of 2 experiments.

memory CD4$^+$ T cells also expressed NKG2D (figure 4G). In contrast, HC vimentin-64-Cit$_{59-71}$-reactive and vimentin$_{59-71}$-reactive CD4$^+$ T cells did not differ in phenotype from total CD4$^+$ T cells (online Supplementary figure 2). Thus, in INA FDRs and patients with RA but not HC, vimentin-64-Cit$_{59-71}$-reactive and vimentin$_{59-71}$-reactive CD4$^+$ T cells reflect antigen-driven activation and differentiation.

TCR bias and oligoclonal TCR reactive with vimentin-64-Cit$_{59-71}$ and vimentin$_{59-71}$

To obtain evidence of in vivo expansion in HLA-DRB1*14:02$^+$ INA, we used multiplex PCR to sequence the TCRs from a total of 53 single vimentin$_{59-71}$-reactive and 71 vimentin-64Cit$_{59-71}$-reactive tetramer-positive CD4$^+$ T cells derived from five patients with HLA-DRB1*14:02$^+$ RA and five HLA-DRB1*14:02$^+$

Table 2 Repeated TCR TRAV, TRBV and TRAV/TRBV clonotypes used by HLA-DRB1*14:02-restricted vimentin$_{59\text{-}71}$ or vimentin-64Cit$_{59\text{-}71}$-reactive CD4$^+$ T cells

Subject	Tetramer	TRAV	CDR3α	TRAJ	Frequency number/Total, (%)	*HLA-DRB1*
RA W1:1	Cit-vimentin	13–2	SQPGTAL	15	2/17 (11.8%)	14:02, 14:02
RA W1:2	Vimentin	26–1	SGAGSYQL	28	5/13 (38%)	04:04, 14:02
Subject	**Tetramer**	**TRBV**	**CDR3β**	**TRBJ**	**Frequency* (%)**	
RA W1:1	Cit-vimentin	2	SEAADNEQ	2–1	2/14 (14.2%)	14:02, 14:02
RA W1:2	Vimentin	10–3	GGTRTESSYEQ	2–7	2/5 (40%)	04:04, 14:02
FDR W1:3*	Cit-vimentin	30	SIGAGNQPQ	1–5	1/19 (5.2%)	04:04, 14:02
FDR W1:3*	Vimentin	30	SIGAGNQPQ	1–5	3/7 (42.8%)	
FDR W1:5†	Cit-vimentin	30	SVGAGNQPQ	1–5	2/2 (100%)	13:02, 14:02
FDR W1:5†	Vimentin	30	SVGAGNQPQ	1–5	2/6 (33%)	
Subject	**Tetramer**	**TRAV/ TRBV**	**CDR3a/CDR3β**	**TRAJ/TRBJ**	**Frequency* (%)**	
RA W1:1	Cit-vimentin	13-2/2	SQPGTAL/ SEAADNEQ	15/2–1	2/12 (16.6%)	14:02, 14:02

Subject	TRBV								
FDR W1:3†	S	I	G	A	G	N	Q	P	Q
	agt	atc	ggg	gcg	ggc	aat	cag	ccc	cag
FDR W1:5†	S	V	G	A	G	N	Q	P	Q
	agt	gtg	ggg	gca	ggc	aat	cag	ccc	cag

TCR repertoire analysis was undertaken in RA patients W1:1–5 and FDR W1:1–5. Productive single TRAV and TRBV clonotypes detected from two patients with RA and three FDRs are shown.

*Frequency reflects the frequency of the repeated TRAV or TRBV clonotype divided by the total number of tetramer$^+$ cells with productive TRAV or TRBV sequence, for each individual.

†Nucleotide sequences encoding each of the public CDR3β amino acid sequences, which require a minimal number of N or P additions to be produced. Nucleotides attributed by the germline Vβ, Dβ and Jβ genes are shown in blue, red and green, respectively. N-additions are in black and P-additions in purple text. The nucleotides at the D–J junction encoding the same amino acid are underlined in each case.

FDR, first-degree relatives; RA, rheumatoid arthritis.

FDRs. *TRAV* and *TRBV* gene usage among CD4$^+$ T cells reactive to each vimentin epitope was generally diverse in RA and FDR (online supplementary table 4). However, *TRBV20-1* and *TRBV30* were preferentially used variable gene segments for recognition of vimentin 64-Arg and 64-Cit variant peptides bound to HLA-DRB1*14:02 (figure 4H). *TRAV13-2* and *TRAV26-1* were preferentially used variable gene segments among both vimentin$_{59\text{-}71}$-reactive and vimentin-64Cit$_{59\text{-}71}$-reactive TCRs. Use of preferential variable gene segments in the repertoires of T cells recognising each epitope was observed in both patients with RA and FDRs and suggested underlying oligoclonality of the autoreactive T cell populations when compared with the total CD3$^+$ T cell population (figure 4H). Indeed, multiple vimentin-64Cit$_{59\text{-}71}$-reactive and vimentin$_{59\text{-}71}$-reactive CD4$^+$ T cells bearing the same CDR3α and/or CDR3β sequences were identified among the single cells sorted from two of the patients with RA and two FDRs (table 2). In two FDRs, CD4$^+$ T cells bearing the same TRBV30 CDR3 sequences were identified multiple times among single cells reactive for vimentin 64-Arg and 64-Cit variant peptides bound to HLA-DRB1*14:02. In one patient with RA, the same TRBV2 CDR3 sequences were identified multiple times among single vimentin-64Cit$_{59\text{-}71}$-reactive CD4$^+$ T cells, and in another patient the same TRBV10-3 CDR3 sequences were identified multiple times among single vimentin$_{59\text{-}71}$-reactive CD4$^+$ T cells. These repeated CDR3α and CDR3β sequences indicate antigen-reactive clonal expansion within the blood of these HLA-DRB1*14:02$^+$ patients and at-risk FDRs (table 2). In all cases, clonally expanded tetramer$^+$ T cells were CD4$^+$CD25$^+$CD127$^+$ Teff, as determined by index sorting. Individuals in whom any oligoclonal sequences were detected in peripheral blood (PB) were more likely to be *HLA-DRB1*14:02* homozygous or *HLA-DRB1*14:02/*04:04* compound heterozygous than individuals without oligoclonal expansion ($p<0.05$, X^2 test). Remarkably, a common TRBV30 CDR3 sequence (SI/VGAGNQPQ) was expanded in the blood of two individual FDRs, which in each case encoded TCRs recognising vimentin$_{59\text{-}71}$ as well as vimentin-64Cit$_{59\text{-}71}$. Of 34 sorted antigen-reactive T cells yielding productive TRBV gene sequences from FDRs, 23.5% contained this CDR3β sequence. We used retroviral vectors encoding HLA-DR14:02-vimentin$_{59\text{-}71}$-restricted TCR P2F3 (online supplementary table 4, bold) to transduce the αβ TCR-deficient SKW-3 cell line.[32 33] When stimulated with HLA-DRB1*14:02-vimentin$_{59\text{-}71}$ or vimentin-64Cit$_{59\text{-}71}$ tetramers, P2F3 SKW-3 cells upregulated phospho- extracellular signal-related kinases (pERK) relative to HLA-DRB1*14:02-CLIP tetramers (figure 4I), confirming that TCR identified from cells with HLA-DRB1*14:02-vimentin$_{59\text{-}71}$ tetramer reactivity recognises vimentin$_{59\text{-}71}$ and vimentin-64Cit$_{59\text{-}71}$ in the context of HLA-DRB1*14:02.

The biased TCR usage suggests a structural requirement for conserved amino acid sequences to recognise vimentin$_{59\text{-}71}$ and vimentin-64Cit$_{59\text{-}71}$. Nucleotide sequences encoding TRBV30 CDR3 reveal that although the second N region at the D–J junction in each TCR is different, they encode the same amino acid sequence (table 2). These data implicate convergent recombination events in the selection of this sequence during TCR gene rearrangement.[34]

DISCUSSION

*HLA-DRB1*14:02* and *HLA-DRB1*04:04* are shown to be independent risk alleles for ACPA$^+$ RA in the INA population. Analysis of the structures and T cell responses to Cit and non-Cit epitopes shows that, consequent to the His13βSer

polymorphism and altered P4 pocket, HLA-DRB1*14:02 can present both variant peptides. This contrasts with structures of HLA-DRB1*04:01 and *04:04, in which Cit but not Arg can be accommodated in P4.[12 29] Presentation of both 64-Cit and 64-Arg vimentin$_{59\text{-}71}$ variants promoted autoreactive CD4$^+$ T cell activation and differentiation to Teff and Treg, and clonal expansion of Teff in patients with *HLA-DRB1*14:02*$^+$ RA and at-risk FDRs. In *HLA-DRB1*14:02*+ HC, we observed no activation or differentiation of antigen-specific T cells above the background total CD4+ T cells. Previous studies of CD4+ T cells in individuals carrying Caucasian SE alleles show that T cell responses to Cit-autoantigenic peptides are increased relative to Arg-variant peptides,[12 29 35 36] reinforcing that T cell function aligns with HLA-DR-peptide structure.

We show preferential variable gene segments and clonally expanded TCR among vimentin$_{59\text{-}71}$-reactive and vimentin-64Cit$_{59\text{-}71}$-reactive CD4$^+$CD25$^+$CD127$^+$ Teff sorted from patients with RA and FDRs, including a public TRBV CDR3. These data suggest that presentation of vimentin self-epitopes in vivo continues in genetically predisposed individuals before and after onset of RA, and selects T cells making productive TCR rearrangements, as identified by the nucleotide sequences, for antigen recognition and T cell expansion.

Although limited by small sample numbers, the phenotypic profiles of 64-Cit and 64-Arg variant-specific autoreactive T cells appeared to be different. The fibroblast antigen, vimentin, has widespread tissue expression.[37] Vimentin 64-Arg-specific memory CD4+ T cells specifically expressed CD69. CD69 is a marker of recent activation or tissue residency and exposure to cytokines such as tumour necrosis factor (TNF),[30] suggesting that vimentin$_{59\text{-}71}$-reactive T cells are activated in tissue inflammatory sites. NKG2D signifies Teff costimulatory function and is TNF-activated.[31] Intriguingly, our data suggest cross-reactivity of some TCRs for the 64-Cit and 64-Arg variants: cell frequency of each specificity was correlated, the public clonotype occurred in T cells reactive with each variant, and T cells expressing TCR P2F3 were activated with both tetramers. In general, coexpansion of T cells recognising Cit and Arg variants might be advantageous, for example, for cross-protection against infectious antigens requiring rapid immunity.[38–40] This suggests a hypothesis for persistence of *HLA-DRB1*14:02* in the INA population, even though its molecular structure permits presentation of multiple self-antigens driving ACPA$^+$ RA.

Low-titre ACPA develops in healthy individuals independent of HLA-DR SE, particularly in inflammatory contexts.[6–8 10 11] In HLA-DR SE$^+$ individuals, memory T cells driven by antigen experience would provide required B cell help for increased titres and epitope spreading in patients developing RA.[8 41 42] Since HLA-DRB1*14:02 broadens capacity for autoantigen presentation and T cell expansion, our study provides a mechanism for enhanced risk of early onset of ACPA+ RA in INA.

Author affiliations
[1]Department of Biochemistry and Molecular Biology, Infection and Immunity Program, Biomedicine Discovery Institute Monash University, Clayton, Australia
[2]The University of Queensland Diamantina Institute, Translational Research Institute, Princess Alexandra Hospital, Brisbane, Australia
[3]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[4]Department of Medicine/Immunology and Rheumatology, Stanford University, VA Palo Alto Health Care System, Palo Alto, California
[5]Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[6]Department of Microbiology and Immunology, The University of Melbourne, Peter Doherty Institute for Infection and Immunity, Melbourne, Australia
[7]Faculty of Agricultural Technology, Technological Educational Institute of Ionian Islands, Argostoli Kefalonia, Greece
[8]Laboratory of Biochemistry and Biophysics, Faculty of Agricultural Technology, Epirus Institute of Technology, Arta, Greece
[9]Arthritis Centre, University of Manitoba, Winnipeg, Manitoba, Canada
[10]Division of Community Health Services, Alaska Native Tribal Health Consortium, Anchorage, Alaska, USA
[11]Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, USA
[12]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, USA
[13]Center for Data Sciences, Brigham and Women's Hospital, Boston, Massachusetts, USA
[14]Arthritis Research UK Centre for Genetics and Genomics, University of Manchester, Manchester, UK
[15]Institute of Infection and Immunity, Cardiff University School of Medicine, Heath Park, Cardiff, UK
[16]Australian Research Council Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton, Australia

Acknowledgements We thank the staff at the National Australian Synchrotron for assistance with data collection, and staff at the Monash Macromolecular Crystallisation Facility and Nishta Ramnoruth for assistance with flow cytometric staining.

Contributors Concept and design of the study: RT, HE-G, JR, RET, SWS, S-CL. Data collection, data analysis and interpretation: SWS, S-CL, YTT, JvH, JS, AJD, EBC, AKM, GKP, DvdW, IS, CAH, DBR, EDF, CNB, XM, VA, TH, KK, HHR, SR. Manuscript preparation: SWS, S-CL, JvH, RT, JR, RET, HE-G. All authors revised and approved the manuscript to be published.

Funding This work was supported by the Australian National Health and Medical Research Council (NHMRC), Australian Research Council (ARC) and Canadian Institutes of Health Research (CIHR: MOP77700) funding. JR is an ARC Australian Laureate Fellow and RT an NHMRC Fellow.

Competing interests There was no commercial support for this work. However we wish to declare that RT is a director of a spin-off company that is commercialising antigen-specific immunotherapy for rheumatoid arthritis.

Ethics approval University of Manitoba, Alaska Area Institutional Review Board, Metro South HREC, University of Queensland HREC.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1 Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. an approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum* 1987;30:1205–13.
2 Raychaudhuri S, Sandor C, Stahl EA, *et al*. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet* 2012;44:291–6.
3 Han B, Diogo D, Eyre S, *et al*. Fine mapping seronegative and seropositive rheumatoid arthritis to shared and distinct HLA alleles by adjusting for the effects of heterogeneity. *Am J Hum Genet* 2014;94:522–32.
4 Ireland JM, Unanue ER. Autophagy in antigen-presenting cells results in presentation of citrullinated peptides to CD4 T cells. *J Exp Med* 2011;208:2625–32.
5 Ireland JM, Unanue ER. Processing of proteins in autophagy vesicles of antigen-presenting cells generates citrullinated peptides recognized by the immune system. *Autophagy* 2012;8:429–30.
6 Padyukov L, Silva C, Stolt P, *et al*. A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. *Arthritis Rheum* 2004;50:3085–92.
7 Quirke AM, Perry E, Cartwright A, *et al*. Bronchiectasis is a Model for chronic bacterial infection inducing autoimmunity in rheumatoid Arthritis. *Arthritis Rheumatol* 2015;67:2335–42.
8 Koning F, Thomas R, Rossjohn J, *et al*. Coeliac disease and rheumatoid arthritis: similar mechanisms, different antigens. *Nat Rev Rheumatol* 2015;11:450–61.
9 Alpizar-Rodriguez D, Brulhart L, Mueller RB, *et al*. The prevalence of anticitrullinated protein antibodies increases with age in healthy individuals at risk for rheumatoid arthritis. *Clin Rheumatol* 2017;36:677–82.
10 Fisher BA, Cartwright AJ, Quirke AM, *et al*. Smoking, Porphyromonas gingivalis and the immune response to citrullinated autoantigens before the clinical onset of rheumatoid arthritis in a southern european nested case-control study. *BMC Musculoskelet Disord* 2015;16:331.

11 Terao C, Asai K, Hashimoto M, *et al*. Significant association of periodontal disease with anti-citrullinated peptide antibody in a japanese healthy population - The Nagahama study. *J Autoimmun* 2015;59:85–90.
12 Scally SW, Petersen J, Law SC, *et al*. A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J Exp Med* 2013;210:2569–82.
13 James EA, Rieck M, Pieper J, *et al*. Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. *Arthritis Rheumatol* 2014;66:1712–22.
14 Peschken CA, Esdaile JM. Rheumatic diseases in North America's indigenous peoples. *Semin Arthritis Rheum* 1999;28:368–91.
15 Ferucci ED, Darrah E, Smolik I, *et al*. Prevalence of anti-peptidylarginine deiminase type 4 antibodies in rheumatoid arthritis and unaffected first-degree relatives in indigenous North American Populations. *J Rheumatol* 2013;40:1523–8.
16 Smolik I, Robinson DB, Bernstein CN, *et al*. First-degree relatives of patients with rheumatoid arthritis exhibit high prevalence of joint symptoms. *J Rheumatol* 2013;40:818–24.
17 Willkens RF, Nepom GT, Marks CR, *et al*. Association of HLA-Dw16 with rheumatoid arthritis in Yakima Indians. further evidence for the "shared epitope" hypothesis. *Arthritis Rheum* 1991;34:43–7.
18 Nelson JL, Boyer G, Templin D, *et al*. HLA antigens in Tlingit Indians with rheumatoid arthritis. *Tissue Antigens* 1992;40:57–63.
19 Williams RC, Jacobsson LT, Knowler WC, *et al*. Meta-analysis reveals association between most common class II haplotype in full-heritage native Americans and rheumatoid arthritis. *Hum Immunol* 1995;42:90–4.
20 Sokolove J, Bromberg R, Deane KD, *et al*. Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. *PLoS One* 2012;7:e35296.
21 Tungatt K, Bianchi V, Crowther MD, *et al*. Antibody stabilization of peptide-MHC multimers reveals functional T cells bearing extremely low-affinity TCRs. *J Immunol* 2015;194:463–74.
22 Wang GC, Dash P, McCullers JA, *et al*. T cell receptor αβ diversity inversely correlates with pathogen-specific antibody levels in human cytomegalovirus infection. *Sci Transl Med* 2012;4:128ra42.
23 Wagner CA, Sokolove J, Lahey LJ, *et al*. Identification of anticitrullinated protein antibody reactivities in a subset of anti-CCP-negative rheumatoid arthritis: association with cigarette smoking and HLA-DRB1 'shared epitope' alleles. *Ann Rheum Dis* 2015;74:579–86.
24 Geluk A, van Meijgaarden KE, Southwood S, *et al*. HLA-DR3 molecules can bind peptides carrying two alternative specific submotifs. *J Immunol* 1994;152:5742–8.
25 Hennecke J, Wiley DC. Structure of a complex of the human alpha/beta T cell receptor (TCR) HA1.7, influenza hemagglutinin peptide, and major histocompatibility complex class II molecule, HLA-DR4 (DRA*0101 and DRB1*0401): insight into TCR cross-restriction and alloreactivity. *J Exp Med* 2002;195:571–81.
26 Stern LJ, Brown JH, Jardetzky TS, *et al*. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 1994;368:215–21.
27 James EA, Moustakas AK, Bui J, *et al*. HLA-DR1001 presents "altered-self" peptides derived from joint-associated proteins by accepting citrulline in three of its binding pockets. *Arthritis Rheum* 2010;62:2909–18.
28 James EA, Moustakas AK, Bui J, *et al*. The binding of antigenic peptides to HLA-DR is influenced by interactions between pocket 6 and pocket 9. *J Immunol* 2009;183:3249–58.
29 Gerstner C, Dubnovitsky A, Sandin C, *et al*. Functional and Structural Characterization of a Novel HLA-DRB1*04:01-Restricted α-Enolase T Cell Epitope in Rheumatoid Arthritis. *Front Immunol* 2016;7:494.
30 Seddiki N, Santner-Nanan B, Martinson J, *et al*. Expression of interleukin (IL)-2 and IL-7 receptors discriminates between human regulatory and activated T cells. *J Exp Med* 2006;203:1693–700.
31 Sakaguchi S. Regulatory T cells: key controllers of immunologic self-tolerance. *Cell* 2000;101:455–8.
32 Gras S, Chen Z, Miles JJ, *et al*. Allelic polymorphism in the T cell receptor and its impact on immune responses. *J Exp Med* 2010;207:1555–67.
33 Beringer DX, Kleijwegt FS, Wiede F, *et al*. T cell receptor reversed polarity recognition of a self-antigen Major histocompatibility complex. *Nat Immunol* 2015;16:1153–61.
34 Kedzierska K, Thomas PG, Venturi V, *et al*. Terminal deoxynucleotidyltransferase is required for the establishment of private virus-specific CD8+ TCR repertoires and facilitates optimal CTL responses. *J Immunol* 2008;181:2556–62.
35 Law SC, Street S, Yu CH, Ch Y, *et al*. T-cell autoreactivity to citrullinated autoantigenic peptides in rheumatoid arthritis patients carrying HLA-DRB1 shared epitope alleles. *Arthritis Res Ther* 2012;14:R118.
36 James EA, Rieck M, Pieper J, *et al*. Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. *Arthritis Rheumatol* 2014;66:1712–22.
37 Kishaba Y, Matsubara D, Niki T. Heterogeneous expression of nestin in myofibroblasts of various human tissues. *Pathol Int* 2010;60:378–85.
38 Furman D, Jojic V, Sharma S, *et al*. Cytomegalovirus infection enhances the immune response to influenza. *Sci Transl Med* 2015;7:281ra43.
39 Chang MH, Moonesinghe R, Athar HM, *et al*. Trends in disparity by sex and race/Ethnicity for the leading causes of death in the United States-1999-2010. *J Public Health Manag Pract* 2016;22 Suppl 1(Suppl 1):S13–S24.
40 Foote EM, Singleton RJ, Holman RC, *et al*. Lower respiratory tract infection hospitalizations among american indian/Alaska native children and the general United States child population. *Int J Circumpolar Health* 2015;74:29256.
41 Hunt L, Hensor EM, Nam J, *et al*. T cell subsets: an immunological biomarker to predict progression to clinical arthritis in ACPA-positive individuals. *Ann Rheum Dis* 2016;75:1884–9.
42 Janssen KM, Westra J, Chalan P, *et al*. Regulatory CD4+ T-Cell subsets and Anti-Citrullinated protein antibody repertoire: potential biomarkers for Arthritis Development in Seropositive Arthralgia Patients? *PLoS One* 2016;11:e0162101.

OPEN ACCESS

EXTENDED REPORT

Dominant B cell receptor clones in peripheral blood predict onset of arthritis in individuals at risk for rheumatoid arthritis

Paul P Tak,[1] Marieke E Doorenspleet,[2,3] Maria J H de Hair,[1] Paul L Klarenbeek,[2,3] Marian H van Beers-Tas,[4] Antoine H C van Kampen,[5] Dirkjan van Schaardenburg,[4] Danielle M Gerlag,[1] Frank Baas,[3] Niek de Vries[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-2017-211351).

[1]Department of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
[2]Amsterdam Rheumatology and Immunology Center, Academic Medical Center, Amsterdam, The Netherlands
[3]Department of Genome Analysis, Academic Medical Center, Amsterdam, The Netherlands
[4]Amsterdam Rheumatology and Immunology Center, Reade, Amsterdam, The Netherlands
[5]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, Amsterdam, The Netherlands

Correspondence to
Niek de Vries, Department of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam, Amsterdam 1100 DE, The Netherlands; n.devries@amc.uva.nl

Received 21 February 2017
Revised 15 July 2017
Accepted 18 July 2017
Published Online First 8 August 2017

To cite: Tak PP, Doorenspleet ME, de Hair MJH, *et al*. *Ann Rheum Dis* 2017;**76**:1924–1930.

ABSTRACT

Background The onset of seropositive rheumatoid arthritis (RA) is preceded by the presence of specific autoantibodies in the absence of synovial inflammation. Only a subset of these *at-risk* individuals will develop clinical disease. This impedes efforts to implement early interventions that may prevent onset of clinically manifest disease. Here we analyse whether clonal changes in the B cell receptor (BCR) repertoire can reliably predict onset of signs and symptoms.

Methods In a prospective cohort study in 21 individuals at risk for RA based on the presence of autoantibodies, the BCR repertoire of paired peripheral blood and synovial tissue samples was analysed using next-generation BCR sequencing. BCR clones that were expanded beyond 0.5% of the total repertoire were labelled dominant. The relative risk (RR) for onset of arthritis was assessed using the presence of ≥5 dominant BCR clones as cut-off. Findings in peripheral blood were validated in an independent prospective cohort of 50 *at-risk* individuals. Based on the test cohort, individuals in the validation cohort were considered positive if peripheral blood at study entry showed ≥5 dominant BCR clones.

Findings Both in the test and validation cohort, the presence of ≥5 dominant BCR clones in peripheral blood was significantly associated with arthritis development after follow-up (validation cohort RR 6.3, 95% CI 2.7 to 15, $p<1\times10^{-4}$). Even when adjusted for a recently described clinical prediction rule the association remained intact (RR 5.0, 95% CI 1.2 to 20, p=0.024). When individuals developed arthritis, dominant BCR clones disappeared from peripheral blood and appeared in synovial tissue, suggesting a direct role of these clones in disease pathogenesis.

Interpretation Dominant BCR clones in peripheral blood predict onset of clinical signs and symptoms of RA in *at-risk* individuals with high accuracy. Our data suggest that during onset of RA these clones shift from peripheral blood to the target tissue.

INTRODUCTION

Rheumatoid arthritis (RA) is a prototypic chronic autoimmune disease with partly unknown aetiology. Clinically manifest arthritis due to synovial inflammation is the hallmark feature of RA. However, it is not the first sign of disease, as patients may already experience arthralgia and the development of synovial inflammation may be preceded by the presence of disease-specific autoantibodies.[1–3] This situation is reminiscent of that in several other immune-mediated inflammatory diseases.[4–7]

RA-specific autoantibodies, IgM-rheumatoid factor (RF) and/or anticitrullinated protein antibodies (ACPA), can be present up to 15 years before onset of disease.[1 8 9] Towards the onset of clinically evident arthritis the ACPA repertoire may broaden due to epitope spreading,[10 11] and levels of inflammatory cytokines and chemokines may increase.[12 13] Although the presence of ACPA is highly specific for RA[14] and may precede its onset, only 20% of the autoantibody positive subjects will develop arthritis within 4 years.[15]

The presence of these autoantibodies preceding the development of RA clearly points to a role for B cells and plasma cells in the pathogenesis of RA. The pathogenic role of B cells in established RA is supported by the known association with autoantibodies,[16] marked infiltration of the synovium by B cells and plasma cells,[17] the production of autoantibodies in the synovial compartment[18] and the response to B cell-depleting therapy.[19] Consistent with this notion, B-cell receptor (BCR) repertoire analysis showed that dominant clones were found in the inflamed synovial tissue of patients with established RA.[20]

We hypothesised that dominant clones may be detected by BCR sequencing in the peripheral blood during the preclinical phase of RA. This might help predict which *at-risk* individuals will develop arthritis over time. We tested this hypothesis analysing paired peripheral blood and synovial tissue samples from individuals *at risk* for developing RA in a prospective cohort study. We found that the presence of dominant peripheral blood BCR clones can predict future onset of RA, and we validated these findings in an independent cohort. Of interest, during the transition to clinically manifest arthritis the BCR clones were not traceable in peripheral blood anymore, but they were found in synovial tissue as highly dominant clones, pointing to a shift of BCR clones to the synovial compartment. The observation that dominant peripheral blood BCR clones can predict future onset of disease may be relevant for other B cell-mediated autoimmune diseases as well.

Table 1 Clinical characteristics of healthy controls, *at-risk* individuals who did not develop arthritis over time and *at-risk* individuals who developed arthritis. *At-risk* individuals have elevated titres for IgM-RF (>12.5 kU/L) and/or anti-CCP (>25 kAU/L). Healthy individuals have low titres for IgM-RF (≤12.5 kU/L) and anti-CCP (≤25 kAU/L)

	Healthy individuals (n=10)	*At-risk* individuals no arthritis developed (n=10)	*At-risk* individuals arthritis developed (n=11)
Female sex, n (%)	7 (70)	5 (50)	7 (64)
Age, years, median (IQR)	34 (28–51)	50 (39–60)	48 (42–54)
IgM-RF positive, n (%)	0 (0)	7 (70)	7 (64)
Level low positive, n (%)†*	–	6 (86)	4 (57)
Level high positive, n (%)†*	–	1 (14)	3 (43)
ACPA positive, n (%)	0 (0)	7 (70)	9 (82)
Level, median (IQR)‡*	–	920 (549–2491)	470 (144–1781)
IgM-RF/ACPA double positive, n (%)	–	4 (40)	5 (46)
ESR, median (IQR)§	–	3 (2–23)	8 (7–15)
CRP, median (IQR)¶	0.9 (0.4–2.9)	2.1 (1.6–6.3)	6.2 (1.5–10.0)
68TJC, median (IQR)	0 (0)	2 (0–7)	4 (1–10)
66SJC, median (IQR)	0 (0)	0 (0)	0 (0)

Positive IgM-RF: >12.5 kU/L.
Positive anti-CCP2 >25 kAU/L.
*Only in individuals who were positive.
†Levels were categorised into high/low positive according to cut-off levels used in the 2010 ACR/EULAR criteria for rheumatoid arthritis because of changed reference values over time.
‡Measured in kAU/L
§Measured in mm/hour.
¶Measured in mg/L.
ACPA, anticyclic citrullinated peptide antibodies (using anti-CCP2 test); CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; IgM-RF, rheumatoid factor of the IgM isotype; 66SJC, swollen joint count assessed in 66 joints; 68TJC, tender joint count assessed in 68 joints.

METHODS

Study subjects

Sixty-five consecutive individuals without arthritis, but *at risk* for the development of RA defined by the presence of IgM-RF and/or ACPA (anti-CCP2 test, Eurodiagnostica), were prospectively followed (further denoted as '*at-risk* individuals').[2 21] From the 65 included individuals, we randomly selected 10 autoantibody positive *at-risk* individuals who did not develop arthritis (median follow-up 69 (range 42–78) months), and 11 individuals who did develop arthritis (median follow-up 15 (range 0–65) months) as test cohort. Nine individuals of the latter group fulfilled the 2010 ACR/EULAR criteria for RA at onset of arthritis,[22 23] while two had unclassified arthritis at the moment of development of arthritis but subsequently did fulfil RA criteria over time. In addition, 10 autoantibody negative healthy individuals without any joint complaints were included as controls (clinical characteristics of all three groups described in table 1 and online supplementary table S1).

In total, 21 *at-risk* individuals and 10 healthy controls were included in this part of the study.

A validation cohort was used consisting of 50 consecutively included individuals with elevated ACPA and/or IgM-RF without any signs of arthritis and at least 36 months follow-up (further details are described in ref [24]). During sequencing and bioinformatic analysis for dominant clones laboratory personnel was blinded for clinical data and outcome.

The cohort studies were approved by the local medical ethical committees of the Academic Medical Center/University of Amsterdam and MC Slotervaart Amsterdam, and all study subjects gave written informed consent.

Peripheral blood and synovial tissue sampling and processing

In the 21 *at-risk* individuals of the test cohort, mini-arthroscopic synovial biopsy sampling was performed upon inclusion in a (non-arthritic) knee joint as previously described.[25] Peripheral blood samples were drawn and stored in PAXGene Blood RNA tubes according to the manufacturer's instructions (catalogue #762165, PreAnalytiX, Breda, the Netherlands). Storage of synovial biopsies, isolation and quantification of RNA, and cDNA synthesis were performed as described previously.[26] Mini-arthroscopy in *at-risk* individuals who subsequently developed arthritis was performed on the same joint, after patients fulfilled the 2010 ACR/EULAR criteria for RA[22 23] and before initiation of treatment.

Linear amplification and next-generation sequencing (NGS)

The linear amplification protocol has been extensively described before.[26] Details are provided in the online supplementary methods. Samples were prepared for next-generation sequencing according to the manual for amplicon sequencing, and sequenced on a Roche Genome Sequencer FLX (Titanium platform). 10,000 BCR_{heavy} sequences were analysed for each peripheral blood sample and 7500 BCR_{heavy} sequences for each synovial tissue sample. We use the term dominant BCR clone to denote clones whose unique BCR signals represent ≥0.5% of the repertoire, as described earlier.[20]

Bioinformatics pipeline and data analysis

The bioinformatics pipeline used to obtain the BCR sequences was described previously in detail[27] and contains four modules: multiplex identifier sorting, identification of V and J gene segments, CDR3 detection and removal of artefacts. Immunoglobulin isotype homology was determined using the National Center for Biotechnology Information's open-access web tool Megablast and reference sequences for the human immunoglobulin heavy-chain constant regions, allowing a sequence homology >97%.[28] Values are expressed as mean and SD or median and IQR range, according to criteria for (non-)parametric analysis. Differences between groups were analysed

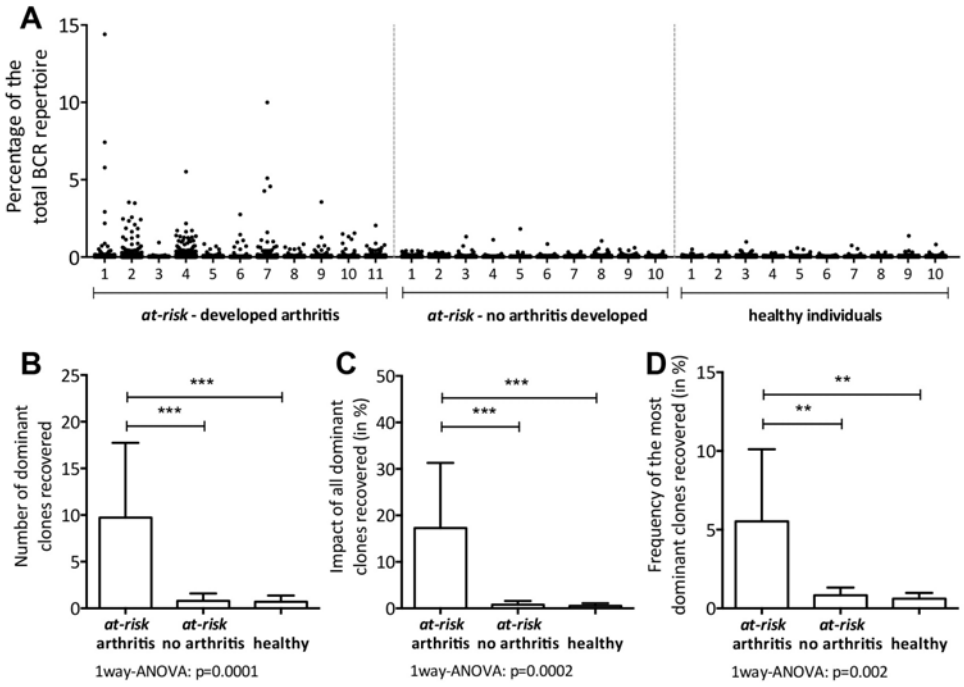


Figure 1 (A) Scatterplot of the BCR repertoire in peripheral blood of 11 *at-risk* individuals who developed arthritis (*at-risk - developed arthritis*), 10 *at-risk* individuals who did not develop arthritis (*at-risk - no arthritis developed*) and 10 autoantibody negative healthy individuals. Each dot represents one clone. The size of the clones is depicted as percentage of the total BCR_{heavy} sequences. (B) The absolute number of dominant BCR clones (clonal size ≥0.5% of the total repertoire), (C) the impact of all dominant clones combined and (D) the size of the single most dominant clone, in *at-risk* individuals who developed arthritis (*at-risk* arthritis, n=11) versus *at-risk* individuals who did not develop arthritis yet (*at-risk* no arthritis, n=10), and healthy individuals (healthy, n=10). Bars show mean and SD, ***p<0.0001, **p<0.001 using one-way ANOVA. ANOVA, analysis of variance; BCR, B-cell receptor.

using Student's t-test, Mann-Whitney U test, one-way analysis of variance or χ^2 test where appropriate. Receiver operating characteristic (ROC) curves were used to determine cut-off values for the prediction of arthritis development in the test cohort. Logistic regression analysis was used to predict the added value of high-throughput fingerprinting and quantitation of BCR clones compared with an existing prediction model for the development of RA.[24] GraphPad Prism software version 6 and PASW Statistics version 22 were used to perform the analyses. p-values <0.05 were considered statistically significant.

RESULTS

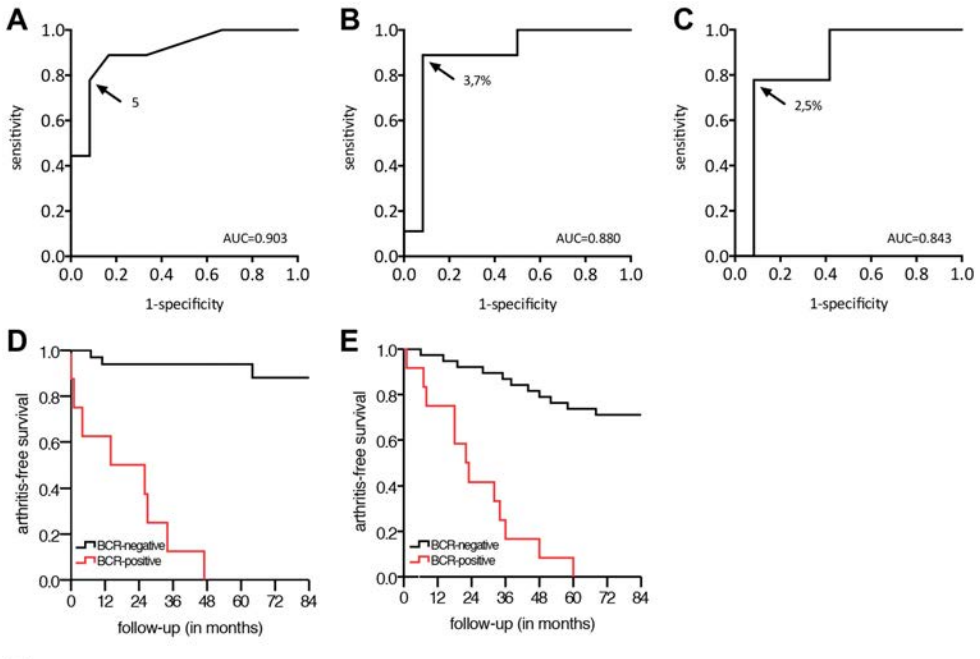
Identification of dominant peripheral blood BCR clones before onset of arthritis

Based on earlier observations that dominant BCR clones are present in the synovial tissue during clinically manifest RA, we hypothesised that these clones might be detectable in the peripheral blood before development of arthritis. Indeed, multiple dominant BCR clones were detected in peripheral blood of all 11 prospectively followed *at-risk* individuals who developed arthritis from the test cohort, as long as 66 months before the clinical onset of arthritis. In contrast, dominant BCR clones were nearly absent both in the 10 *at-risk* individuals who did not develop arthritis and in the 10 healthy individuals (figure 1A). We observed that the number of dominant BCR clones, the impact of all dominant BCR clones combined (size of all dominant clones combined as percentage of the total repertoire) and the impact of the most dominant BCR clone were increased in *at-risk* individuals who developed arthritis during follow-up, compared with those who did not develop arthritis and healthy individuals (number of dominant clones mean 9.7±8.0 vs 0.8±0.8 vs 0.7±0.7, respectively, p=0.001 (figure 1B), impact of the dominant clones median 16.4% of the total repertoire, IQR 3.7%–33.7% vs 0.7% IQR 0%–1.7% vs 0.5% IQR 0%–1.1%, respectively, p<0.0001 (figure 1C) and impact of the single most dominant clone mean 5.5%±4.6% vs 0.7%±0.7% vs 0.6%±0.4%, respectively, p<0.0003 (figure 1D)). Subsequently, we analysed synovial tissue biopsies in *at-risk* individuals obtained during the preclinical phase, but these samples contained BCR mRNA levels that were too low to allow next-generation sequenting (NGS). The low BCR mRNA levels in the synovium during the preclinical stage of the disease are explained by the absence of B cell infiltration as demonstrated before using immunohistochemistry.[3]

Collectively, these observations demonstrate that dominant BCR clones are readily detectable in peripheral blood during the preclinical phase in all *at-risk* individuals who will develop RA after follow-up, but not in those who did not.

The presence of dominant BCR clones predicts future arthritis development

Having shown that the presence of BCR clones can be detected in peripheral blood in all *at-risk* individuals who will subsequently develop RA, in some cases after several years, we next aimed to develop a biomarker that can be used to identify individuals who have a high risk of developing arthritis in the short term. Such patients might be treated in the at-risk phase to prevent onset of arthritis.[29] A clinically relevant follow-up period of 36 months was chosen to evaluate arthritis development. This time period may carry a risk high enough to justify preventive pharmacological intervention, while being short enough to infer urgency for treatment.



F

	test cohort (n=65)	validation cohort (n=50)
Sensitivity	78% (40-96%)	67% (39-87%)
Specificity	92% (60-100%)	94% (79-99%)
PPV	72% (33-94%)	83% (51-97%)
NPV	94% (64-100%)	87% (71-95%)

Figure 2 Receiver operating characteristic curves for (A) the number of dominant clones, (B) the impact of all dominant clones combined and (C) the impact of the most dominant clone in *at-risk* individuals (n=21). The development of arthritis was analysed after 36 months of follow-up. The arrow points to the cut-off value chosen, and the corresponding value is shown. (D) Kaplan-Meier curve for *BCR-clone positive* and *BCR-clone negative* individuals in the test cohort, assuming the *at-risk* individuals analysed represent a random selection of the total *at-risk* individuals (n=65). (E) Kaplan-Meier curve for *BCR-clone positive* and *BCR-clone negative* individuals in the validation cohort. (F) Table describing sensitivity, specificity, PPV and NPV including 95% CIs for the *BCR-clone model*, in the test cohort and the validation cohort. AUC, area under the curve; BCR, B-cell receptor; NPV, negative predictive value; PPV, positive predictive value.

We designed three tests based on the number of dominant BCR clones present, the impact of all dominant clones combined on the BCR repertoire and the impact of the single most dominant BCR clone; ROC curves are depicted in figure 2A–C. Based on these ROC curves, optimal cut-offs were determined at ≥5 dominant BCR clones in peripheral blood, a combined impact ≥3.7% and an impact of the most dominant clone ≥2.5%, respectively. We decided to use the presence of ≥5 dominant BCR clones as comprehensible and intuitive marker for further studies. This is further denoted as '*BCR-clone positive,*' and <5 dominant BCR clones as '*BCR-clone negative,*' and collectively as the *BCR-clone model*.

The cut-off of ≥5 dominant clones in peripheral blood resulted in two clearly distinguishable groups, and corresponding sensitivity of 78%, specificity of 92%, positive predictive value (PPV) of 72% and negative predictive value (NPV) of 94% (figure 2D, 2F for the Kaplan-Meier curve, log rank test p<0.001). We had access to a second independent cohort of 50 subjects to validate our findings using the same cut-off value. Fifteen *at-risk* individuals developed arthritis within 36 months; the characteristics are described in table 2.

Analysis in this validation cohort showed that *BCR-clone positive at-risk* individuals had an 83% risk of developing RA within 36 months (PPV), while this risk was 13% in *at-risk BCR-clone negative* individuals (1-NPV), resulting in a relative risk of 6.3 (95% CI 2.7 to 15, p<0.0001, log rank test p<0.001, figure 2E). Post hoc analysis on both cohorts revealed that within 60 months, all *BCR-clone positive* individuals developed arthritis (after 47, 48 and 60 months, respectively, online supplementary figure S1).

The 50 *at-risk* individuals in the validation cohort were previously used to develop a prediction model for the development of RA,[24] the *risk rule* model. This describes a composite score of multiple clinical parameters categorising *at-risk* individuals into low, intermediate and high-risk individuals (respectively 17, 20 and 13 individuals). Using logistic regression analysis to calculate the added value of the *BCR-clone model* to the existing *risk rule*, an overall relative risk of 5.0 (95% CI 1.2 to 20, p=0.024) was found. In the low, intermediate and high-risk groups the relative risk contributed by BCR-clone positivity was estimated at 18 (95% CI 0.6 to 520), 6.1 (95% CI 1.9 to 20) and 1.2 (95% CI 0.6 to 2.7), respectively.

In conclusion, we show that *at-risk* individuals with five or more dominant BCR clones in peripheral blood have an 83% risk of developing arthritis within 36 months, compared with a risk of 13% in individuals with four or less dominant BCR

Table 2 Clinical characteristics of *at-risk* individuals in the validation cohort

	At-risk individuals no arthritis developed (n=35)	*At-risk* individuals arthritis developed (n=15)
Female sex, n (%)	22 (63)	8 (53)
Age, years, median (IQR)	51 (43–55)	47 (37–52)
IgM-RF positive, n (%)	28 (80)	11 (73)
Level low positive, n (%)†*	20 (57)	7 (64)
Level high positive, n (%)†*	15 (43)	4 (36)
ACPA positive, n (%)	20 (57)	14 (93)
Level, median (IQR)‡*	1333 (364–9650)	342 (155–1016)
IgM-RF/ACPA double positive, n (%)	13 (37)	10 (67)
ESR, median (IQR)§	12 (4–18)	10 (3–19)
CRP, median (IQR)¶	2.4 (0.9–4.5)	2.3 (1.1–9.4)
53TJC, median (IQR)	0 (0–0)	0 (0–1)
Risk rule model**		
Low risk, n (%)	17 (49%)	0 (0%)
Intermediate risk, n (%)	14 (40%)	6 (40%)
High risk, n (%)	4 (11%)	9 (60%)

Positive IgM-RF: >12.5 kU/L.
Positive anti-CCP2 >25 kAU/L.
*Only in individuals who were positive.
†Levels were categorised into high/low positive according to cut-off levels used in the 2010 ACR/EULAR criteria for rheumatoid arthritis because of changed reference values over time.
‡Measured in kAU/L.
§Measured in mm/hour.
¶Measured in mg/L.
**Score based on the *risk rule*[24] scaled 0–13 points.
ACPA, anticyclic citrullinated peptide antibodies (using anti-CCP2 test); CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; IgM-RF, rheumatoid factor of the IgM isotype; 53TJC, tender joint count assessed in 53 joints.

clones. Moreover, analysis after 5 years revealed that all individuals who initially tested positive developed arthritis.

Dominant BCR clones present in peripheral blood during the preclinical phase have migrated to synovial tissue in clinically manifest RA

We hypothesised that if the observed dominant BCR clones are involved in synovial inflammation, then these clones might also be detectable in synovial tissue after onset of RA. To this end, we analysed peripheral blood samples obtained during the preclinical stage, and paired blood and synovial tissue after onset of arthritis in eight individuals who developed arthritis.

Up to 29.3% (median, IQR 9.9%–45.8%) of all preclinically detectable dominant peripheral blood BCR clones were detectable in synovial tissue after development of arthritis. All ranked within the top 25 most dominant clones in synovial tissue (figure 3A–C). Most strikingly, none of the preclinically detectable dominant peripheral blood BCR clones could be recovered from peripheral blood after arthritis developed (figure 3B,C). Additional analyses of the dominant clones found in both peripheral blood and synovial tissue showed that these clones are class switched to the IgG1 isotype and enriched for IGHV4-34, features that are all associated with autoreactivity, while IGHV3-23, IGHV1-69, CDR3 length, ACPA or IgM-RF titres, and HLA-DRB1 shared epitope positivity (SE) are comparable between the three groups (described in online supplementary results, supplementary figures 2 and 3, and supplementary table 2).[30–33]

Together, these analyses show that dominant BCR peripheral blood clones present during the preclinical phase of arthritis are in part retrievable as dominant clones in synovial tissue once arthritis becomes clinically apparent. At this time point, the clones are no longer found in peripheral blood anymore. These migratory clones have features that have been associated with autoreactivity.

DISCUSSION

The results presented here show that the presence of dominant BCR clones in peripheral blood predicts with high accuracy the onset of arthritis in patients who are *at risk* of developing RA. Moreover, we found support that these dominant clones may migrate to synovial tissue once arthritis becomes apparent. These findings are consistent with the notion that B-cell abnormalities occur up to several years before the onset of synovial inflammation, and that the development of RA is a multistep process. Conceivably, a 'second hit,' for instance a trauma or viral infection, may lead to synovial inflammation, subsequent migration of autoreactive B-cell clones towards the synovium and impaired resolution of inflammation in patients with pre-existing systemic autoimmunity.[2 34] This work provides the rationale for future studies on B-cell abnormalities during the preclinical stage in other immune-mediated inflammatory diseases like systemic lupus erythematosus (SLE), multiple sclerosis (MS) and type 1 diabetes (T1D), and opens up the perspective of preventive intervention.

Since not all individuals with RA-specific antibodies progress to clinically manifest RA, the relation between RA-specific antibodies and clonal expansions is unclear. The current data stress once more that the antibodies in the preclinical phase are produced by plasmablasts and long-lived plasma cells located elsewhere (eg, bone marrow and spleen),[35] and B cells and plasmablasts present in blood represent migrating cell populations. Moreover, better biomarkers are needed to predict which *at-risk* individuals will develop RA. Autoantibody and cytokine profiles, specific gene signatures, body mass index, current smoking and autonomic nervous system dysfunction all contribute to the risk of developing RA.[36–41] Our data provide a novel biomarker that has superior predictive power compared with other available biomarkers evaluated so far. It increases the accuracy of the previously reported prediction rule for the development of arthritis in autoantibody positive subjects.[24]

Identification of *at-risk* individuals who will develop RA in the short term enables development of early preventive strategies.[29 42] Our findings support the rationale for B cells or the interaction between B cells and T cells as targets for preventive therapy. The cut-off used here (five or more dominant clones) was chosen to be able to identify subjects with a high risk of developing RA with an acceptable NPV to avoid unnecessary treatment. Whether a preventive pharmacological intervention will be considered acceptable is of course dependent on the benefit/risk profile and the cost-effectiveness of the specific treatment.

As discussed above, there is strong evidence for a role of B cells and plasma cells in the pathogenesis of RA. The development of RA is associated with defects in central and peripheral tolerance leading to autoreactive B cells,[43] and circulating autoantibodies can be detected years before the onset of the disease. It is tempting to speculate that the clones identified in the present study are pathogenic B cells since (1) they are not detected in healthy controls nor in subjects at risk who do not develop RA after follow-up, (2) their dominance suggests

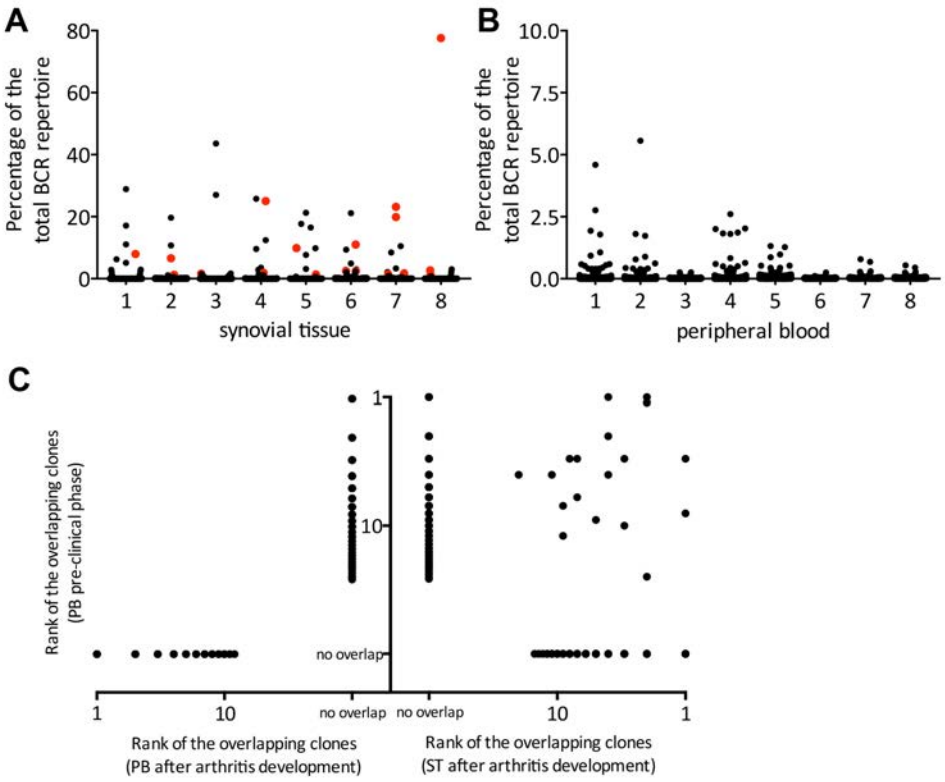


Figure 3 Scatterplot of the BCR repertoire in synovial tissue (A) and peripheral blood (B) after arthritis development in eight patients. Each dot represents one clone. Clones in red represent clones that were dominantly present in peripheral blood during the preclinical phase. The size of the clones is depicted as percentage of the total BCR_{heavy} sequences. (C) Dot plot showing the overlap between dominant BCR clones in the preclinical phase and after arthritis development (n=8). The y-axis depicts the rank of the clones found in blood during the preclinical phase (all eight patients pooled). On the left x-axis the overlap with dominant clones in peripheral blood after arthritis development, on the right x-axis the overlap with dominant clones in synovial tissue after arthritis development. In case no overlap was found, the dots were marked '*no overlap.*' BCR, B-cell receptor; PB, peripheral blood; ST, synovial tissue.

activity, (3) the clones seem to migrate to the inflamed synovial tissue after arthritis development, and (4) these clones show additional features associated with autoreactivity.

There are technical limitations to the study; first, we measure the BCR_{heavy} chain repertoire on mRNA level since it only identifies expressed BCRs and limits the number of amplifications. However, we cannot distinguish whether clonal signatures are derived from memory B cells or plasma cells. Nevertheless, the presence of dominant BCR clones in the preclinical phase of RA is a robust and reproducible marker. Second, to analyse mRNA, cells were lysed preventing further phenotypic characterisation of dominant clones. Further unravelling the phenotype of BCR clones in the preclinical phase is essential to better understand the role of these cells in the earliest phase of disease, and might lead to even more specific biomarkers. Third, coupling of BCR heavy and light chains is prevented by the technique used, thus limiting determination of reactivity. This should be addressed in a future study.

All patients had clinically manifest arthritis at the time of the second synovial biopsy but the joint that was biopsied was not clinically inflamed except for one patient. Still, these biopsies showed a diverse repertoire resembling the repertoires of clinically inflamed joints, containing the dominant predictive BCR clones identified in peripheral blood in the preclinical phase. This can be explained by the fact that clinically uninvolved joints of patients with established RA exhibit histological signs of synovial inflammation, as described before.[44]

Taken together, we show the presence of increased BCR clonal signatures in peripheral blood obtained during the preclinical stage of RA. During onset of arthritis, these BCR clones disappear from blood and appear in the target tissue, where they may drive autonomous disease progression. Our observations show that the presence of dominant BCR clones in peripheral blood in the at-risk stage accurately predicts short-term onset of clinically manifest disease. Consequently, they may serve as a biomarker that could help guide decisions about pharmacological treatment to prevent the onset of clinically manifest disease. This is important since recent studies clearly indicate that early intervention may be more effective and lead more often to drug-free remission.[45] There are already studies that focus on intervention during the pre-RA phase. An example is the recently completed "PRAIRI" study (http://www.trialregister.nl/trialreg/ admin/rctview.asp?TC=1969), in which individuals at risk of developing RA were treated with a single course of rituximab to delay development of clinically manifest arthritis. The current marker determined during the preclinical phase can be used to further investigate the effect of therapeutical intervention on the clonal distribution over time. Similar studies are currently under way with abatacept and simvastatin. Future work should explore whether BCR clones might also predict onset of disease in other immune-mediated inflammatory disorders like type 1 diabetes, multiple sclerosis, SLE and vasculitis.

Acknowledgements We would like to thank Rebecca Esveldt, Barbera van Schaik, Marja Jakobs, Ted Bradley and Dr Lisa van Baarsen for technical assistance. This work was carried out on the Dutch national e-infrastructure with the support of SURF Foundation. Parts of this work have been presented at a scientific meeting and/or published as a conference abstract.

Contributors All authors contributed to the formulation of the hypotheses and research questions and the analysis plan, and provided critical input into the draft manuscript. PPT, MJHdH, MHvBT, DVS and DMG were involved in recruitment of patients to the study. MED and PLK performed the experiments under supervision of PPT, FB and NdV. MED, MJHdH, PLK, AHCvK and NdV did the statistical analysis. PPT, MED and NdV drafted the manuscript.

Funding This study was partially supported by the Dutch Arthritis Association (grant no 06-1-303). MED, NdV, DMG and PPT received support from CTMM, the Center for Translational Molecular Medicine, project TRACER (grant 04I-202), and BTCURE, a research project from the Innovative Medicines Initiative Joint Undertaking (grant no 115142-2). PLK was supported by an AMC Graduate School PhD Scholarship.

Competing interests None declared.

Ethics approval Academic Medical Center/University of Amsterdam and MC Slotervaart Amsterdam, The Netherlands.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Nielen MM, van Schaardenburg D, Reesink HW, *et al*. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum* 2004;50:380–6.
2 van de Sande MG, de Hair MJ, van der Leij C, *et al*. Different stages of rheumatoid arthritis: features of the synovium in the preclinical phase. *Ann Rheum Dis* 2011;70:772–7.
3 de Hair MJ, van de Sande MG, Ramwadhdoebe TH, *et al*. Features of the synovium of individuals at risk of developing Rheumatoid arthritis: implications for understanding preclinical rheumatoid arthritis. *Arthritis Rheumatol* 2014;66:513–22.
4 Berger T, Rubner P, Schautzer F, *et al*. Antimyelin antibodies as a predictor of clinically definite multiple sclerosis after a first demyelinating event. *N Engl J Med* 2003;349:139–45.
5 Wasserfall CH, Atkinson MA. Autoantibody markers for the diagnosis and prediction of type 1 diabetes. *Autoimmun Rev* 2006;5:424–8.
6 Eriksson C, Kokkonen H, Johansson M, *et al*. Autoantibodies predate the onset of systemic lupus erythematosus in northern Sweden. *Arthritis Res Ther* 2011;13:R30.
7 van Schaik FD, Oldenburg B, Hart AR, *et al*. Serological markers predict inflammatory bowel disease years before the diagnosis. *Gut* 2013;62:683–8.
8 Aho K, Heliövaara M, Maatela J, *et al*. Rheumatoid factors antedating clinical rheumatoid arthritis. *J Rheumatol* 1991;18:1282–4.
9 Rantapää-Dahlqvist S, de Jong BA, Berglin E, *et al*. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum* 2003;48:2741–9.
10 van der Woude D, Rantapää-Dahlqvist S, Ioan-Facsinay A, *et al*. Epitope spreading of the anti-citrullinated protein antibody response occurs before disease onset and is associated with the disease course of early arthritis. *Ann Rheum Dis* 2010;69:1554–61.
11 van de Stadt LA, van der Horst AR, de Koning MH, *et al*. The extent of the anti-citrullinated protein antibody repertoire is associated with arthritis development in patients with seropositive arthralgia. *Ann Rheum Dis* 2011;70:128–33.
12 Deane KD, O'Donnell CI, Hueber W, *et al*. The number of elevated cytokines and chemokines in preclinical seropositive rheumatoid arthritis predicts time to diagnosis in an age-dependent manner. *Arthritis Rheum* 2010;62:3161–72.
13 Kokkonen H, Söderström I, Rocklöv J, *et al*. Up-regulation of cytokines and chemokines predates the onset of rheumatoid arthritis. *Arthritis Rheum* 2010;62:383–91.
14 Schellekens GA, Visser H, de Jong BA, *et al*. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum* 2000;43:155–63.
15 Bos WH, Wolbink GJ, Boers M, *et al*. Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study. *Ann Rheum Dis* 2010;69:490–4.
16 Schellekens GA, de Jong BA, van den Hoogen FH, *et al*. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. *J Clin Invest* 1998;101:273–81.
17 Tak PP, Smeets TJ, Daha MR, *et al*. Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. *Arthritis Rheum* 1997;40:217–25.
18 Vossenaar ER, Smeets TJ, Kraan MC, *et al*. The presence of citrullinated proteins is not specific for rheumatoid synovial tissue. *Arthritis Rheum* 2004;50:3485–94.
19 Edwards JC, Szczepanski L, Szechinski J, *et al*. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. *N Engl J Med* 2004;350:2572–81.
20 Doorenspleet ME, Klarenbeek PL, de Hair MJ, *et al*. Rheumatoid arthritis synovial tissue harbours dominant B-cell and plasma-cell clones associated with autoreactivity. *Ann Rheum Dis* 2014;73:756–62.
21 Gerlag DM, Raza K, van Baarsen LG, *et al*. EULAR recommendations for terminology and research in individuals at risk of Rheumatoid arthritis: report from the study group for risk factors for rheumatoid arthritis. *Ann Rheum Dis* 2012;71:638–41.
22 Aletaha D, Neogi T, Silman AJ, *et al*. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.
23 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.
24 van de Stadt LA, Witte BI, Bos WH, *et al*. A prediction rule for the development of arthritis in seropositive arthralgia patients. *Ann Rheum Dis* 2013;72:1920–6.
25 van de Sande MG, Gerlag DM, Lodde BM, *et al*. Evaluating antirheumatic treatments using synovial biopsy: a recommendation for standardisation to be used in clinical trials. *Ann Rheum Dis* 2011;70:423–7.
26 Klarenbeek PL, de Hair MJ, Doorenspleet ME, *et al*. Inflamed target tissue provides a specific niche for highly expanded T-cell clones in early human autoimmune disease. *Ann Rheum Dis* 2012;71:1088–93.
27 Klarenbeek PL, Tak PP, van Schaik BD, *et al*. Human T-cell memory consists mainly of unexpanded clones. *Immunol Lett* 2010;133:42–8.
28 Giudicelli V, Chaume D, Lefranc MP. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res* 2005;33:D256–61.
29 Gerlag DM, Norris JM, Tak PP. Towards prevention of autoantibody-positive rheumatoid arthritis: from lifestyle modification to preventive treatment. *Rheumatology* 2016;55:607–14.
30 Schröder AE, Greiner A, Seyfert C, *et al*. Differentiation of B cells in the nonlymphoid tissue of the synovial membrane of patients with rheumatoid arthritis. *Proc Natl Acad Sci U S A* 1996;93:221–5.
31 Pugh-Bernard AE, Silverman GJ, Cappione AJ, *et al*. Regulation of inherently autoreactive VH4-34 B cells in the maintenance of human B cell tolerance. *J Clin Invest* 2001;108:1061–70.
32 Wardemann H, Yurasov S, Schaefer A, *et al*. Predominant autoantibody production by early human B cell precursors. *Science* 2003;301:1374–7.
33 Cambridge G, Moura RA, Santos T, *et al*. Expression of the inherently autoreactive idiotope 9G4 on autoantibodies to citrullinated peptides and on rheumatoid factors in patients with early and established rheumatoid arthritis. *PLoS One* 2014;9:e107513.
34 Buckley CD, Gilroy DW, Serhan CN, *et al*. The resolution of inflammation. *Nat Rev Immunol* 2013;13:59–66.
35 Manz RA, Thiel A, Radbruch A. Lifetime of plasma cells in the bone marrow. *Nature* 1997;388:133–4.
36 van Baarsen LG, Bos WH, Rustenburg F, *et al*. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. *Arthritis Rheum* 2010;62:694–704.
37 Sokolove J, Bromberg R, Deane KD, *et al*. Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. *PLoS One* 2012;7:e35296.
38 Lübbers J, Vosslamber S, van de Stadt LA, *et al*. B cell signature contributes to the prediction of RA development in patients with arthralgia. *Ann Rheum Dis* 2015;74:1786–8.
39 de Hair MJ, Landewé RB, van de Sande MG, *et al*. Smoking and overweight determine the likelihood of developing rheumatoid arthritis. *Ann Rheum Dis* 2013;72:1654–8.
40 Koopman FA, Tang MW, Vermeij J, *et al*. Autonomic dysfunction precedes development of Rheumatoid arthritis: a prospective cohort study. *EBioMedicine* 2016;6:231–7.
41 Moura RA, Cascão R, Perpétuo I, *et al*. Cytokine pattern in very early rheumatoid arthritis favours B-cell activation and survival. *Rheumatology* 2011;50:278–82.
42 Tak PP. Are we ready to change the pace of arthritis treatment? treating pre-arthritis and very early arthritis. *Acta Reumatol Port* 2011;36:8–9.
43 Samuels J, Ng YS, Coupillaud C, *et al*. Impaired early B cell tolerance in patients with rheumatoid arthritis. *J Exp Med* 2005;201:1659–67.
44 Kraan MC, Versendaal H, Jonker M, *et al*. Asymptomatic synovitis precedes clinically manifest arthritis. *Arthritis Rheum* 1998;41:1481–8.
45 Raza K, Saber TP, Kvien TK, *et al*. Timing the therapeutic window of opportunity in early rheumatoid arthritis: proposal for definitions of disease duration in clinical trials. *Ann Rheum Dis* 2012;71:1921–3.

EXTENDED REPORT

Pan-PPAR agonist IVA337 is effective in experimental lung fibrosis and pulmonary hypertension

Jerome Avouac,[1,2] Irena Konstantinova,[3] Christophe Guignabert,[4,5] Sonia Pezet,[1] Jeremy Sadoine,[6] Thomas Guilbert,[1] Anne Cauvet,[1] Ly Tu,[4,5] Jean-Michel Luccarini,[3] Jean-Louis Junien,[3] Pierre Broqua,[3] Yannick Allanore[1,2,6]

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-2016-210821).

[1]Université Paris Descartes, Sorbonne Paris Cité, INSERM U1016, Institut Cochin, CNRS UMR8104, Paris, France
[2]Université Paris Descartes, Sorbonne Paris Cité , Service de Rhumatologie A, Hôpital Cochin, Paris, France
[3]Inventiva, Daix, France
[4]INSERM UMR_S 999, Le Plessis-Robinson, France
[5]Université Paris-Sud, Université Paris-Saclay, Le Kremlin-Bicêtre, France
[6]EA 2496 Pathologie, Imagerieet Biothérapies Orofaciales, UFR Odontologie, Université ParisDescartes and PIDV, PRES Sorbonne Paris Cité, Montrouge, France

Correspondence to
Dr Jerome Avouac, Service de Rhumatologie A, Hôpital Cochin, Université Paris Descartes, 27 rue du Faubourg Saint-Jacques, 75014 Paris, France; javouac@me.com

Received 13 November 2016
Revised 7 July 2017
Accepted 18 July 2017
Published Online First 11 August 2017

To cite: Avouac J, Konstantinova I, Guignabert C, *et al*. *Ann Rheum Dis* 2017;**76**:1931–1940.

ABSTRACT

Objective To evaluate the antifibrotic effects of the pan-peroxisome proliferator-activated receptor (PPAR) agonist IVA337 in preclinical mouse models of pulmonary fibrosis and related pulmonary hypertension (PH).

Methods IVA337 has been evaluated in the mouse model of bleomycin-induced pulmonary fibrosis and in Fra-2 transgenic mice, this latter being characterised by non-specific interstitial pneumonia and severe vascular remodelling of pulmonary arteries leading to PH. Mice received two doses of IVA337 (30 mg/kg or 100 mg/kg) or vehicle administered by daily oral gavage up to 4 weeks.

Results IVA337 demonstrated at a dose of 100 mg/kg a marked protection from the development of lung fibrosis in both mouse models compared with mice receiving 30 mg/kg of IVA337 or vehicle. Histological score was markedly reduced by 61% in the bleomycin model and by 50% in Fra-2 transgenic mice, and total lung hydroxyproline concentrations decreased by 28% and 48%, respectively, as compared with vehicle-treated mice. IVA337 at 100 mg/kg also significantly decreased levels of fibrogenic markers in lesional lungs of both mouse models. In addition, IVA337 substantially alleviated PH in Fra-2 transgenic mice by improving haemodynamic measurements and vascular remodelling. In primary human lung fibroblasts, IVA337 inhibited in a dose-dependent manner fibroblast to myofibroblasts transition induced by TGF-β and fibroblast proliferation mediated by PDGF.

Conclusion We demonstrate that treatment with 100 mg/kg IVA337 prevents lung fibrosis in two complementary animal models and substantially attenuates PH in the Fra-2 mouse model. These findings confirm that the pan-PPAR agonist IVA337 is an appealing therapeutic candidate for these cardiopulmonary involvements.

INTRODUCTION

Fibrotic diseases impose a major socioeconomic burden on modern societies and account for up to 45% of deaths in the developed world.[1] The identification of key factors that drive fibrosis are of interest for clinical therapy because, to date, very few drugs have been approved, and they have limited efficacy in preventing progression or reverting existing fibrosis. Fibrosis occurs as a result of sustained injury to the epithelium, which causes the overproduction of cytokines and growth factors. These latter promote the recruitment and differentiation of mesenchymal cell precursors into myofibroblasts, which produce high amounts of collagen and other extracellular matrix proteins.

Nuclear receptors are a family of transcription factors with key roles in fibrotic responses.[2] Peroxisome proliferator-activated receptors (PPARs) are nuclear receptors, the activation of which is known to display antifibrotic and anti-inflammatory properties[3–7]: PPARα activators prevent lung fibrosis, PPARδ agonists reduce bleomycin-induced inflammation and PPARγ agonists attenuate skin, lung and vascular fibrosis.[3 8]

IVA337 is a new chemical entity that activates the three PPAR isoforms. The antifibrotic properties of this product have been assessed in several in vitro and in vivo preclinical studies: IVA337 has been shown to prevent and induce the regression of pre-existing fibrotic damage in the liver and in the skin.[3 9] These preclinical promising results together with the good safety profile of this product in phase I and phase IIa studies have led to further clinical development, investigating the efficacy of IVA337 on skin fibrosis in patients with early diffuse systemic sclerosis (SSc) (NCT02503644). SSc is a life-threatening connective tissue disease of autoimmune origin, considered as a prototype entity for fibrotic diseases. SSc is characterised by pathological fibrosis of the skin and internal organs (2). Pulmonary fibrosis, a common complication of SSc, is associated with substantial mortality and has no approved therapy.[10] Pulmonary hypertension related to SSc (SSc-PH) is also associated with high morbidity and mortality, as well as poorer response to therapy and worse outcomes compared with the idiopathic form of PAH (IPAH). Moreover, the current therapies of SSc-PH or IPAH remain essentially palliative and do not reverse the progressive remodelling of the pulmonary vasculature, which causes increased pulmonary artery pressures and fatal heart failure.[11–13]

Thus, given the very severe prognosis and the lack of efficient treatment of pulmonary fibrosis and PH, our objective was to evaluate the potential efficacy of IVA337 in two preclinical mouse models of these complications: the bleomycin-induced lung fibrosis mouse model and the Fra-2 transgenic mouse model.

MATERIAL AND METHODS

An extended method section is available in the online data supplement.

Effects of IVA337 in bleomycin-induced lung fibrosis and Fra-2 transgenic mice

Mice were treated with oral gavage once a day with vehicle or IVA337 (30 mg/kg or 100 mg/kg) for 15 days in the bleomycin mouse model and 4 weeks in the Fra-2 mouse model.

Assessment of fibrosing alveolitis

The severity of fibrosing alveolitis was assessed in both mouse models according to semiquantitative histological analyses performed on paraffin-embedded lung sections stained with H&E and to the measurement of collagen content in lesional lung samples by the hydroxyproline assay.[14 15]

In the Fra-2 mouse model, lung fibrosis was also assessed by micro-CT and non-linear microscopy with second harmonic generation processing, as previously reported.[16]

Lung biomarker measurement

Selected fibrogenic markers were quantified by real-time PCR or ELISA in the lesional lungs of bleomycin-treated mice and Fra-2 transgenic mice.

Haemodynamic measurements and assessment of pulmonary vascular changes in Fra-2 transgenic mice

Right ventricular systolic pressure (RVSP) and heart rate were determined in unventilated mice under isoflurane anaesthesia. Right ventricular hypertrophy (RVH) was determined by the Fulton index measurement.[17–19] Morphometric analyses were performed on paraffin-embedded lung sections stained using H&E and alpha smooth muscle actin (α-SMA).[20]

Immunostaining of lesional lung sections

The expression of PPARα, PPARδ and PPARγ was assessed by immunofluorescence using appropriate antibodies. Myofibroblast quantification was performed by immunohistochemistry for α-SMA. The numbers of infiltrating T cells, B cells and macrophages were quantified by immunohistochemistry using antibodies targeting CD3, CD22 and CD68, respectively. The nuclear accumulation of phosphorylated smad2/smad3 (pSmad2/3) detected by immunofluorescence was used to reflect the activation of TGF-β signalling.

Lung fibroblast proliferation and activation

The proliferation of human primary pulmonary fibroblasts (HPF cells) on Platelet-derived growth factor (PDGF) stimulation was assessed by the measurement of 5-ethynyl-2′-deoxyuridine (EdU) incorporation. TGF-β-induced fibroblast to myofibroblast transition was assessed by immunostaining with α-SMA.

Statistical analysis

All data are expressed as mean values±SEM. Multiple group comparisons were analysed by using post hoc Dunnett's test. Unpaired or paired t-test was used for a two-group comparison. $p<0.05$ (all two sided) was considered significant.

RESULTS

Tolerance to IVA337

Treatment with IVA337 was well tolerated in both mouse models with no weight loss during the whole treatment period (vehicle: +0.69±1.27 g, IVA337 30 mg/kg: +0.51±1.15 g and IVA337 100 mg/kg: +1.41±1.17 g) and a clinical score of welfare (ranging from 0 to 3) not significantly different between the different groups (vehicle: 0.50±0.63, IVA337 30 mg/kg: 0.12±0.35 and IVA337 100 mg/kg: 0.25±0.38).

IVA337 prevents bleomycin-induced lung fibrosis

IVA337 demonstrated a marked protection from the development of lung fibrosis induced by bleomycin comparatively to vehicle-treated mice. Indeed, IVA337 100 mg/kg strongly reduced by 61% tissue density on histological measurements ($p<0.01$) when compared with vehicle-treated mice (figure 1A and B). Consistent with histological analysis, IVA337 100 mg/kg reduced total lung hydroxyproline concentrations by 28% ($p<0.05$) and myofibroblast counts by 60% ($p<0.05$), as compared with vehicle (figure 1C and D andonline supplementary file). IVA337 100 mg/kg also significantly decreased mRNA levels of *Col1a1*, *Col3a1* (all $p<0.001$) and *Fn1* (all $p<0.05$) in lesional lungs (online supplementary figure S2A–C).

IVA337 alleviates lung fibrosis in the Fra-2 mouse model

We next tested the efficacy of IVA337 in the Fra-2 mouse model, characterised by the spontaneous development of a progressive non-specific interstitial pneumonia. At week 17, Fra-2 mice treated with IVA337 100 mg/kg displayed a significant 21% decrease in lung density (in Hounsfield units (HU)) as compared with Fra-2 mice receiving the vehicle (−524.4 vs −432.2 HU, $p<0.05$) when assessed by chest micro-CT imaging (figure 2A and B). Consistent with this finding, functional residual capacity increased significantly by 30% in mice treated with IVA337 100 mg/kg (72.8% vs 54.1%, $p<0.05$) (figure 2C).

Lung specimens from Fra-2 mice treated with vehicle exhibited features of non-specific interstitial pneumonia (33) (figure 3A). On treatment with IVA337, a significant 58% reduction of the lung fibrosis score was observed at a dose of 100 mg/kg compared with mice treated with the vehicle (figure 3A and B). Consistent with CT and histological analysis, hydroxyproline content and myofibroblast counts were also reduced by 54% ($p<0.05$) and 48% ($p<0.05$), respectively (figure 3C and D). In addition, mRNA levels of *Col1a1*, *Col1a2* and *Fn1*were decreased by IVA337 (online supplementary figure S2D–F).

Second harmonic generation showed in vehicle-treated mice a preferential perivascular distribution of fibrosis, which was consistent with fibrosing alveolitis (figure 3E). Scoring of fibrillar collagen deposits confirmed a significant decrease in collagen scoring in Fra-2 mice receiving IVA337 100 mg/kg, as compared with Fra-2 mice treated with the vehicle or with IVA 30 mg/kg (figure 3F).

IVA337 reverses PH in the Fra-2 mouse model

On treatment with IVA337 100 mg/kg, a substantial reduction in values of RVSP (29.1±1.4 mm Hg vs 34.3±1.3 mm Hg, $p<0.05$) and RVH (0.29±0.01 vs 0.34±0.01 arbitrary units (AU), $p<0.01$) was observed compared with vehicle-treated mice (figure 4A and B). Consistent with these findings, IVA337 100 mg/kg was associated with significant decrease in percent medial wall thickness (figure 4C and E) and numbers of muscularised distal pulmonary arteries (figure 4D and F).

IVA337 decreases the levels of fibrogenic markers in lesional lungs

Successful targeting of the TGF-β signalling axis was observed on treatment with IVA337 in both mouse models. In the bleomycin model, treatment with IVA337 100 mg/kg led to a marked reduction of *Tgfb1* ($p<0.01$), *Tgfb2* ($p<0.05$) and *Tgfbr1* ($p<0.01$)

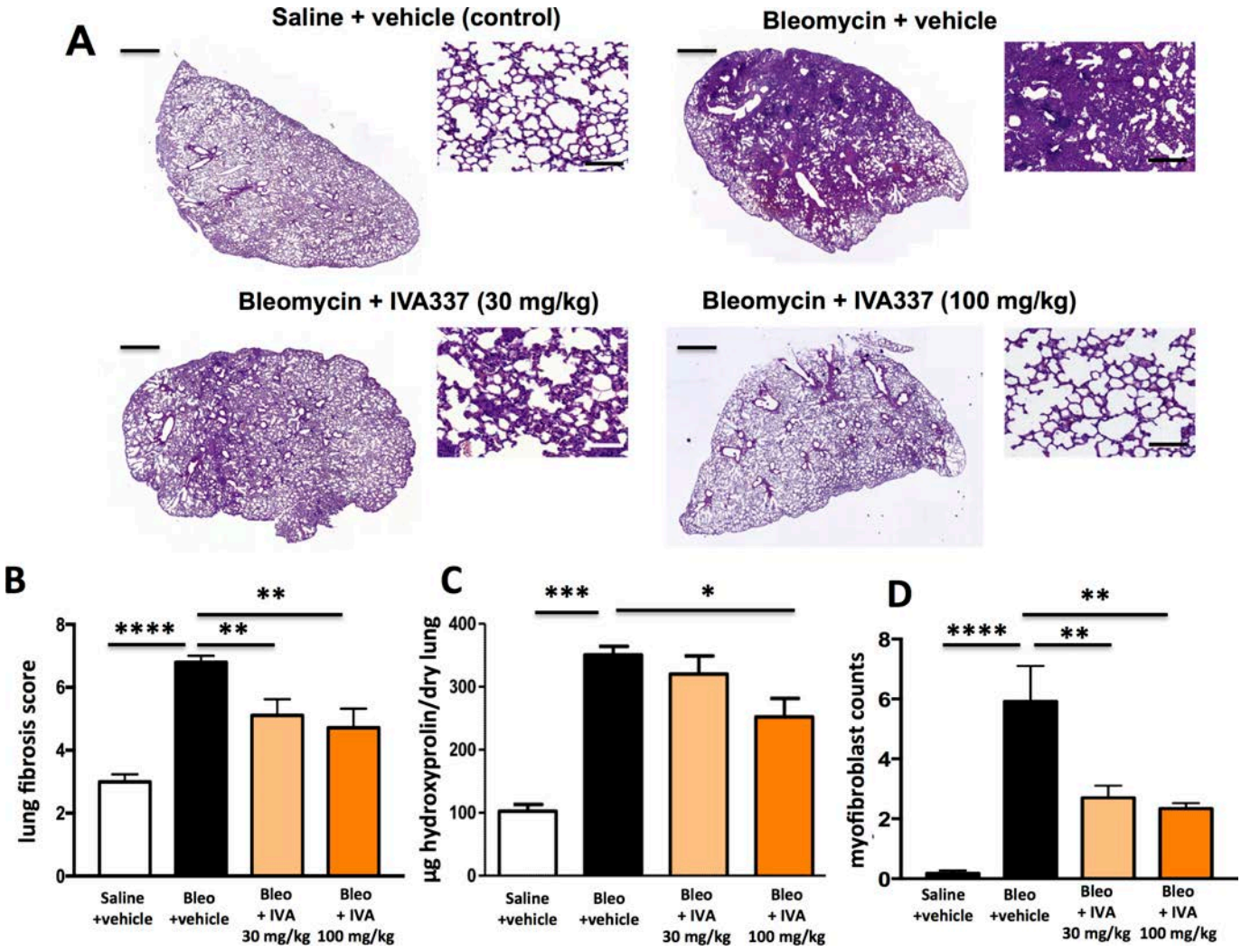


Figure 1 IVA337 100 mg/kg prevents bleomycin-induced lung fibrosis. (A) Lung fibrosis was significantly higher in mice treated with the vehicle or IVA337 30 mg/kg than in those receiving IVA337 100 mg/kg. Representative lung sections stained by H&E are shown. Scale bar=2000 µm and 100 µm (higher magnification). (B) Fibrosis lung histological score was significantly decreased in mice treated with IVA337 100 mg/kg compared with those receiving IVA 30 mg/kg or vehicle-treated mice. (C) Hydroxyproline content was significantly reduced in mice treated with IVA337 100 mg/kg compared with mice receiving IVA337 30 mg/kg or vehicle-treated mice. (D) Myofibroblast counts were also markedly decreased on treatment with IVA337. A total of 44 mice were used (11/group). Values are the mean±SEM. Statistical analyses: one-way ANOVA followed by Dunnett's multiple comparison test. *p<0.05; **p<0.01. ANOVA, analysis of variance.

mRNA levels (online supplementary figure S3A–C), which was not significant at the TGF-β-protein level (online supplementary figure S3D). A significant decrease of nuclear levels of phosphorylated Smad2/Smad3 (pSmad2/3) compared with vehicle-treated mice was also observed on both doses of IVA337 (online supplementary figure S3E and F).

In the Fra-2 model, reduced TGF-β protein levels were detected in lesional lungs (online supplementary figure S4A) and a significant decrease of nuclear levels of phosphorylated Smad2/Smad3 (pSmad2/3) was observed compared with vehicle-treated mice (online supplementary figure S4B).

A striking reduction of TIMP1 protein levels was also detected on treatment with IVA337 compared with vehicle-treated mice in the bleomycin model (figure 5A) and in Fra-2 transgenic mice (figure 5B). Levels of osteopontin (OPN) and monocyte chimoattractant protein-1 (MCP1) were also markedly reduced on treatment with IVA337 100 mg/kg in Fra-2 mice, but not in the bleomycin model (figure 5C–F).

IVA337 reduces T cell, B cell and macrophage infiltration in lesional lungs

To analyse whether treatment with IVA337 influences the outcome of both mouse models by regulating inflammatory infiltrates, we next counted the number of T cells, B cells and macrophages in lesional lungs. T cell, B cell and macrophage counts detected by immunohistochemistry for CD3, CD22 and CD68, respectively, were markedly reduced on IVA337 100 mg/kg in both mouse models (online supplementary figures S5A–D and S6A–D).

IVA337 restores PPARs expression in lesional lungs

To assess a role of IVA337 in regulating the expression levels of PPARs, we examined PPAR expression in both mouse models. The expression of the three PPAR isoforms was markedly decreased after bleomycin challenge. Pan-PPAR activation by IVA337 led to the restoration of PPARα, PPARδ and PPARγ expressions (figure 6A and B). Consistently, IVA337 at 100 mg/kg also markedly restored PPARα and PPARγ expression in lesional lungs of

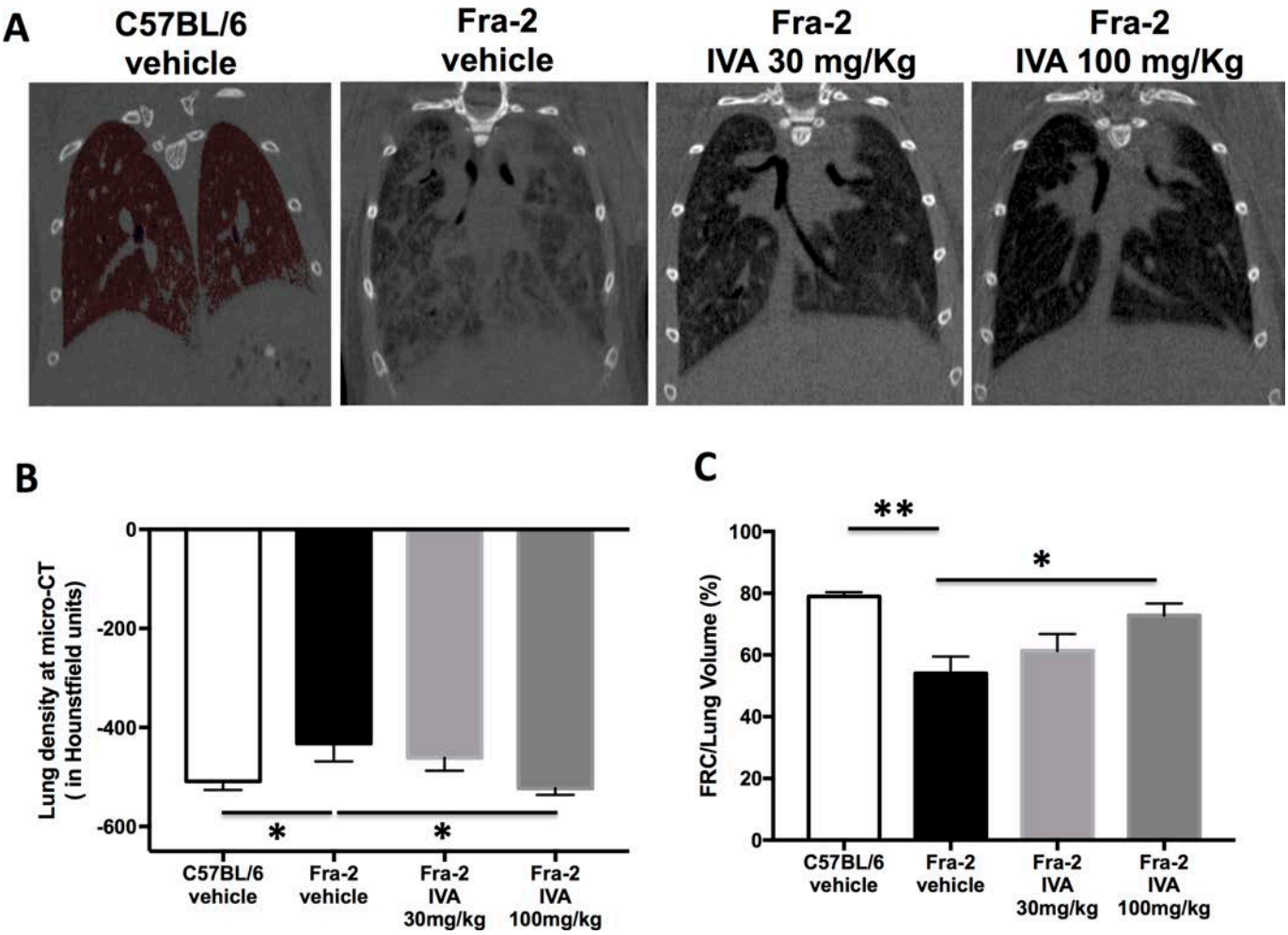


Figure 2 IVA-337 100 mg/kg prevents lung fibrosis in Fra-2 transgenic mice: evaluation by CT-scan. (A) Treatment with IVA337 100 mg/kg prevents lung fibrosis in Fra-2 transgenic mice: representative pictures of micro-CT. (B) Decreased lung density at micro-CT in Fra-2 transgenic mice treated with IVA337 100 mg/kg compared with vehicle-treated mice. (C) Reduced residual lung volume, expressed as the percentage of functional residual capacity on total lung volume, in Fra-2 transgenic mice treated with IVA337 100 mg/kg compared with vehicle-treated mice. A total of 28 mice were used (five C57BL6 mice, seven Fra-2 vehicle, eight Fra-2 IVA337 30 mg/kg and eight Fra-2 IVA337 100 mg/kg). Values are the mean±SEM. Statistics: one-way ANOVA followed by Dunnett's multiple comparison test (comparison of the three groups of Fra-2 mice) or unpaired t test (comparison of the two groups of vehicle-treated mice). *p<0.05; **p<0.01. ANOVA; analysis of variance.

Fra-2 transgenic mice (figure 6C and D). A trend was observed for increased PPARδ expression on treatment with 100 mg/kg IVA337 in Fra-2 mice lungs (figure 6D).

IVA337 inhibits lung human fibroblast proliferation and differentiation

We next aimed to determine whether antifibrotic effects of IVA337 might be due to reduced lung fibroblast proliferation induced by PDGF and/or activation induced by TGF-β. As expected, PDGF-induced proliferation of primary HPF as indicated by an increase of EdU-positive cells (online supplementary figure S7A). PDGF-induced proliferation of primary HPF was markedly inhibited by IVA337, as indicated by a dose-dependent decrease of EdU-positive cells (online supplementary figure S7B).

TGF-β induced fibroblast to myofibroblast transdifferentiation (FMT) in primary HPF as indicated by an increase of SMA-positive cells (measured by immunofluorescence) (online supplementary figure S7C). FMT induced by TGF-β was efficiently inhibited by IVA337, as indicated by a dose-dependent decrease of SMA-positive cells (online supplementary figure S7D).

IVA337 engages PPARγ in primary human lung fibroblasts

In HPF, the antifibrotic and antiproliferative effects of IVA337 are mainly due to PPARγ activity. To evidence PPARγ target engagement, we performed a loss of function experiment using a siRNA approach in primary HPF (online supplementary figure S8A). The knockdown of PPARγ in primary HPF resulted in a potentiated proliferation in response to PDGF (online supplementary figure S8B). In cells transfected with PPARγ siRNA, the efficacy of IVA337 in inhibiting proliferation was markedly reduced in comparison with the control cells (online supplementary figure S8C), which supports that antiproliferative effects of IVA337 are mediated by PPARγ engagement.

DISCUSSION

Our results highlight the substantial interest of activating PPARs to prevent severe organ damages and fibrosis characterising SSc, in addition to the beneficial effects previously observed in experimental skin fibrosis (3). Indeed, we demonstrate that treatment with IVA337 100 mg/kg reduces lung fibrosis in two complementary animal models and substantially attenuates PH in the Fra-2 mouse model. An originality of this study is the assessment of

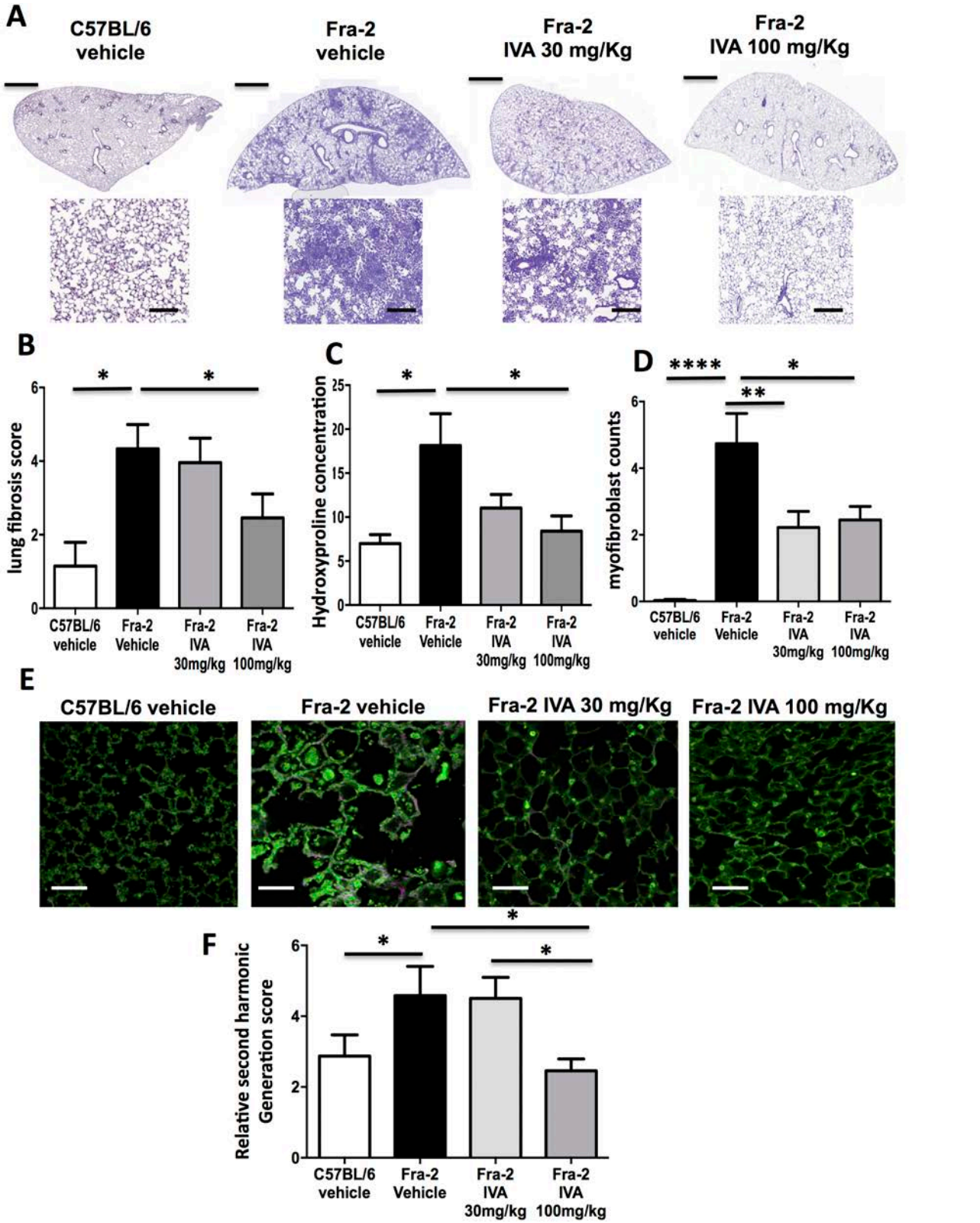


Figure 3 IVA-337 100 mg/kg prevents lung fibrosis in Fra-2 transgenic mice: evaluation by histology. (A) Treatment with IVA337 100 mg/kg prevents lung fibrosis in Fra-2 transgenic mice: representative lung sections stained by H&E. Scale bar: 2000 µm and 100 µm (higher magnification). (B) Histological lung fibrosis score decreased significantly on treatment with IVA337 100 mg/kg, as compared with mice receiving IVA337 30 mg/kg and vehicle-treated mice. (C) Hydroxyproline content in lesional lungs of Fra-2 mice markedly decreased on treatment with IVA337 100 mg/kg compared with vehicle-treated mice. (D) Decreased number of myofibroblasts in lesional lungs on treatment with IVA337. (E) Second harmonic generation showed fibrillar collagen in Fra-2 mice treated with vehicle (in pink), but not in mice receiving IVA337 100 mg/kg. Scale bar: 50 µm. (F) Second harmonic scores were higher in Fra-2 mice receiving the vehicle or IVA337 30 mg/kg, as compared with Fra-2 mice treated with IVA337 100 mg/kg. A total of 28 mice were used (five C57BL6 mice, seven vehicle, eight IVA337 30 mg/kg and eight IVA337 100 mg/kg). Values are the mean±SEM. Statistics: one-way ANOVA followed by Dunnett's multiple comparison test. *p<0.05; **p<0.01; ****p<0.0001. ANOVA, analysis of variance.

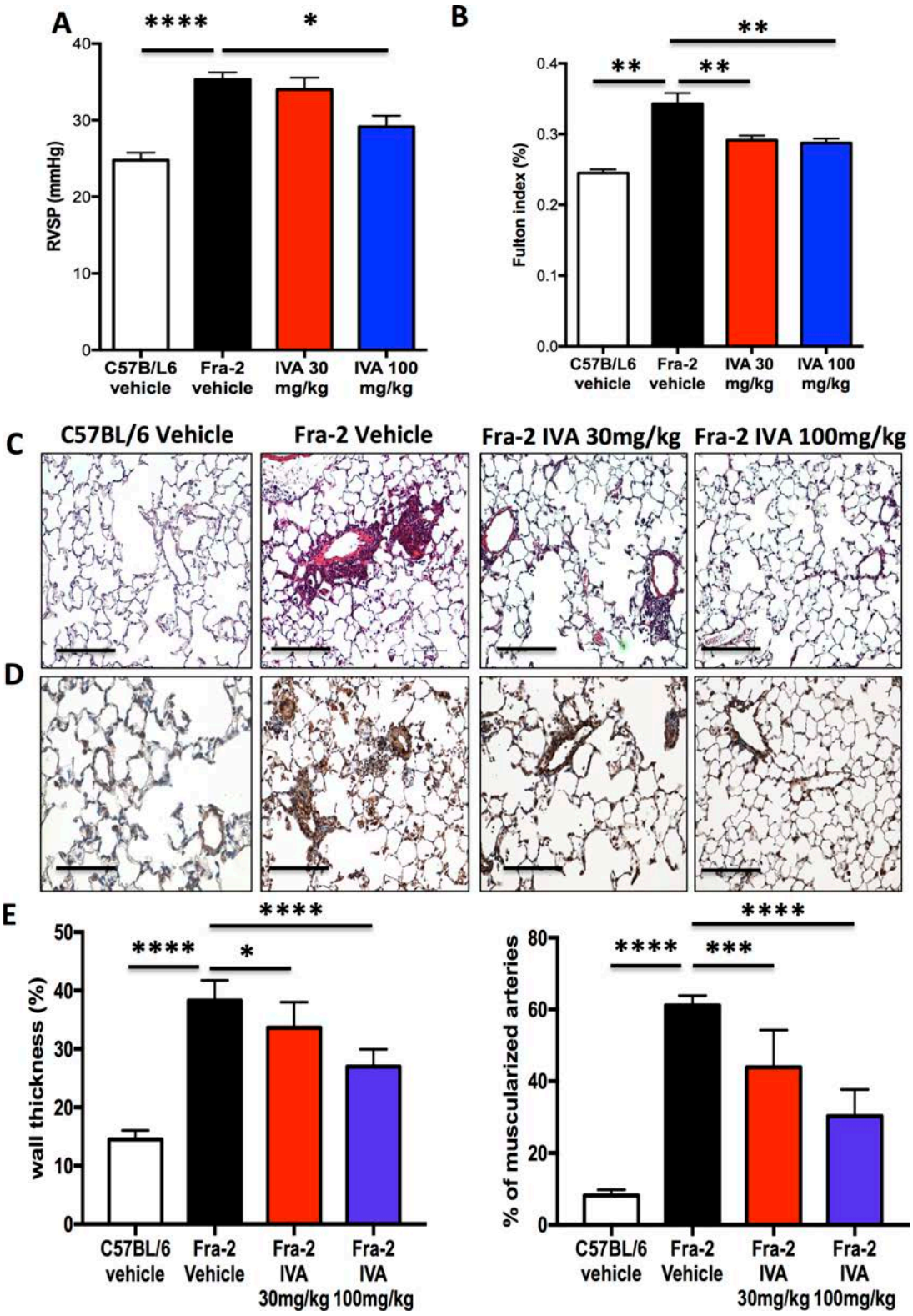


Figure 4 IVA337 100 mg/kg alleviates PH in Fra-2 transgenic mice. (A and B) Right ventricular systolic pressure (RVSP) (A) and right ventricular hypertrophy assessed by the Fulton index (B). A total of 28 mice were used (five C57BL6 mice, seven Fra-2 vehicle, eight Fra-2 IVA337 30 mg/kg and eight IVA337 100 mg/kg). Values are the mean±SEM. Statistics: one-way ANOVA followed by Dunnett's multiple comparison test (comparison of the three groups of Fra-2 mice) or unpaired t-test (comparison of the two groups of vehicle-treated mice). *p<0.05; **p<0.01; ****p<0.0001. (C–F) Representative images of H&E staining (C) showing substantial reduction in the percentage medial wall thickness in Fra-2 mice treated with IVA337 100 mg/kg when compared with Fra-2 mice receiving IVA337 30 mg/kg or the vehicle (E). Representative images of α-smooth muscle actin (α-SMA) immunohistostaining (D) showing significant reduction in the percentage of distal artery muscularisation in lungs of Fra-2 mice treated with IVA337 100 mg/kg when compared with Fra-2 mice receiving IVA337 30 mg/kg or the vehicle (F). Scale bar=100 µm. A total of 28 mice were used (five C57BL6 mice, seven vehicle, eight IVA337 30 mg/kg and eight IVA337 100 mg/kg). Statistical analyses: one-way ANOVA followed by Dunnett's multiple comparison test. *p<0.05; ***p<0.001; ****p<0.0001. ANOVA, analysis of variance.

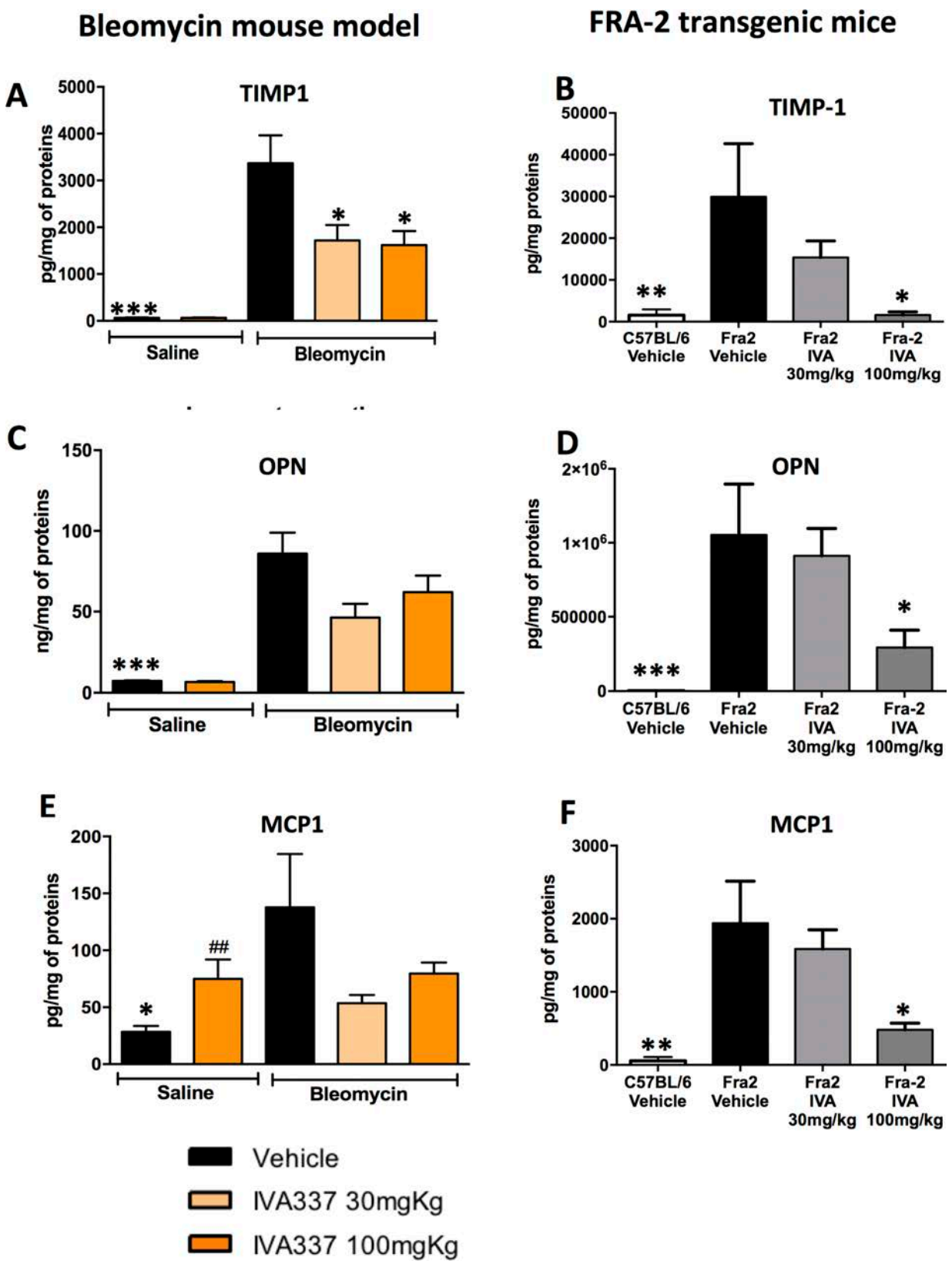


Figure 5 IVA337 100 mg/kg decreases levels of fibrogenic markers in lesional lungs of mice challenged with bleomycin and Fra-2 transgenic mice. (A and B) Protein levels of TIMP1 were markedly decreased on treatment with IVA337 100 mg/kg, as compared with vehicle-treated mice in lesional lungs of mice challenged with bleomycin (A) and Fra-2 transgenic mice (B). (C–F) Protein levels of osteopontin (C) and MCP1 (G) were markedly decreased on treatment with IVA337 100 mg/kg in Fra-2 transgenic mice, as compared with vehicle-treated mice. A trend was observed for decreased concentrations of osteopontin (D) and MCP1 (F) in the bleomycin model. Twenty-three Fra-2 mice and 38 C57BL/6 mice were used for these experiments. Values are represented by dot blots with means±SEM. Statistics: one-way ANOVA followed by Dunnett's multiple comparison test. *p<0.05; **p<0.01; ***p<0.001 versus vehicle (bleomycin vehicle for B and Fra-2 vehicle for D). ANOVA, analysis of variance; TIMP1, TIMP metallopeptidase inhibitor-1.

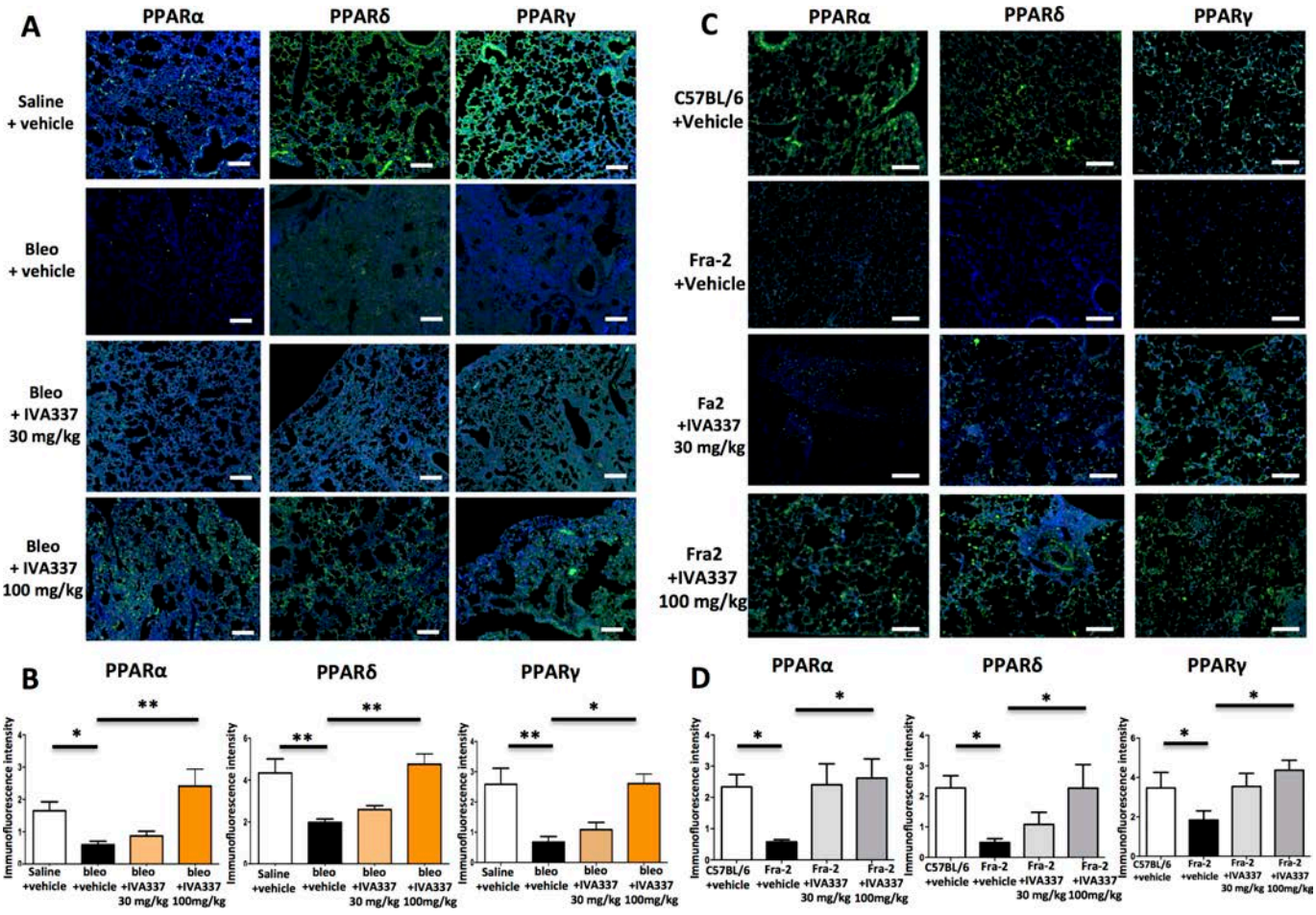


Figure 6 IVA337 engages the different PPAR isoforms in lesional lungs of mice challenged with bleomycin and Fra-2 transgenic mice. (A) PPAR isoform expression in lesional lungs on treatment with IVA337 30 mg/kg, 100 mg/kg or vehicle in the bleomycin mouse model. Representative lung sections stained for PPARα, PPARδ and PPARγ. Scale bar: 100 µm. (B) The expression of the three PPAR isoforms was markedly decreased after bleomycin challenge. IVA337 100 mg/kg led to the activation of its molecular targets in this model, leading to the restoration of PPARα, PPARδ and PPARγ expressions, with a similar level as saline control mice. (C) PPAR isoform expression in lesional lungs on treatment with IVA337 30 mg/kg, 100 mg/kg or vehicle in Fra-2 transgenic mice. Representative lung sections stained for PPARα, PPARδ and PPARγ. Scale bar: 100 µm. (D) IVA337 restored PPARα and PPARγ expressions in Fra-2 transgenic mice. A trend was also observed for increased PPARδ expression on treatment with IVA337. Twenty-three Fra-2 mice (seven vehicle, eight IVA337 30 mg/kg and eight IVA337 100 mg/kg) and 49 C57BL/6 mice (controls of Fra-2 mice: one group of five mice, bleomycin model, 11/group) were used for these experiments. Values are represented by dot blots with means±SEM. Statistics: one-way ANOVA followed by Dunnett's multiple comparison test. *p<0.05; **p<0.01. ANOVA, analysis of variance; PPAR, peroxisome proliferator-activated receptor.

pulmonary fibrosis by micro-CT and PH by right heart catheterisation, two procedures routinely used in Human pathology.

Despite the positive signal observed with IVA337 30 mg/kg at the molecular level, a significant improvement of pulmonary interstitial and vascular diseases was reached only with the dose of 100 mg/kg in both animal models. This result differs from what was observed in dermal fibrosis, in which the effects of IVA337 30 mg/kg and 100 mg/kg were similar. The activation level of different molecular targets may explain this result. Indeed, IVA337 100 mg/kg substantially restored PPARα, PPARδ and PPARγ expression in lesional lung sections of both mouse models, whereas IVA337 at 30 mg/kg only led to a mild effect on the expression of the three PPAR isoforms.

The tolerance of the two dose regimens was similar, and no substantial changes were observed. This emphasises the good safety profile of this drug at 100 mg/kg in the preclinical setting in mice.

Treatment with IVA337 markedly prevents the development of pulmonary fibrosis in the bleomycin mouse model, extending the findings obtained with the PPARα or PPARγ specific agonists.[5 21 22] The advantage to target several PPAR isoforms has been previously suggested by the study of concomitant administration of fenofibrate (PPARα agonist) and rosiglitazone (PPARγ agonist), which enhanced the beneficial effects produced by either fenofibrate or rosiglitazone alone on bleomycin-induced lung fibrosis.[5] Concomitant administration of low doses of fenofibrate and rosiglitazone also provided synergistic renoprotective effect against the development of diabetes-induced nephropathy and fibrosis.[23] IVA337 also displayed potent antifibrotic effects in the Fra-2 mouse model, which is complementary to the bleomycin model since it adds vascular remodelling to inflammation-driven lung fibrosis.[24] Interestingly, Fra-2 directly binds to the PPARγ2 promoter and represses PPARγ2 expression.[25]

The antifibrotic properties of IVA337 are, at least partly, related to a reduction of inflammatory infiltrates, as it was recently shown in inflammation-driven experimental skin fibrosis.[3 5 26] These data are consistent with the regulation by PPARγ ligands of inflammation associated with acute lung disease, including

decreasing release of chemokine/cytokine by alveolar macrophages and neutrophils as well as decreasing migration of these inflammatory cells.[27]

In addition to suppression of the inflammatory response, the attenuation of fibrosis in these models on IVA337 could be attributable to the direct antifibrotic effects of this product. Indeed, IVA337 restores the expression of PPAR isoforms in lesional lung fibroblasts from mice challenged with bleomycin and Fra-2 transgenic mice (online supplementary figure S9) and reduces TGF-β-induced canonical and non-canonical cascades in human fibroblasts.[3] IVA337 also inhibits TGF-β induced collagen synthesis in dermal fibroblasts[3] and directly interferes with primary HPF, the effector cells of pulmonary fibrosis. In the present study, IVA337 inhibited HPF proliferation induced by PDGF in a concentration dependent manner. The knockdown of PPARγ decreased the inhibitory effects of IVA337 on cell proliferation, which supports the hypothesis that this effect is dependent on PPARγ. The antiproliferative effects of PPARγ ligand have been previously demonstrated in some cell types. Ward *et al* showed that the PPARγ ligands 15d-PGJ2 and rosiglitazone could inhibit the proliferation of human cultured airway smooth muscle cells.[28] Treatment with rosiglitazone induced a dose-dependent inhibition of lung adenocarcinoma cells (A549) growth, which was predominantly due to the inhibition of cell proliferation.[29] In addition, IVA337 reduced the α-SMA expression in HPF cells induced by TGF-β. This finding is consistent with the reduction of the TGF-β-induced α-SMA expression observed with IVA337 in dermal fibroblasts[3] and with rosiglitazone in MRC-5 cells.[30] Moreover, potent attenuation of TGF-β-induced collagen protein production has been observed on treatment with PPARγ agonists in human lung fibroblasts.[31]

IVA337 decreases the levels of fibrogenic markers in lesional lungs. In both mouse models, IVA337 markedly reduced the activation of TGF-β signalling. In addition, decreased levels of OPN, a fibrogenic cytokine that promotes migration, adhesion and proliferation of fibroblasts in the development of lung fibrosis,[32] and TIMP1, a key factor to fibrogenic response,[33] were observed on treatment with IVA337 in the bleomycin model and in Fra-2 mice.

PH remains a devastating condition, particularly in patients with SSc. Despite advances in medical therapies, PH continues to cause significant morbidity and mortality, highlighting the need for progress in the identification and validation of potential new targets for therapeutic development against this life-threatening disease. PPARs and particularly PPARγ are expressed in the lung and pulmonary vasculature, and PPARγ expression is reduced in the vascular lesions of patients with PH.[34] Furthermore, it has been demonstrated that the disruption of PPARγ signalling in endothelial cell in mice is sufficient to cause mild PH and to impair recovery from chronic hypoxia-induced PH.[17] In addition, PPARγ is also reduced in the vascular lesions of rats with severe PH caused by treatment with hypobaric hypoxia and a vascular endothelial growth factor receptor antagonist. In our study, pan-PPAR activation mediated by IVA337 alleviated PH in Fra-2 transgenic mice, with a significant improvement of signs of PH (RVSP and RVH), vascular remodelling and myointimal proliferation. Our findings are in agreement with the previous studies supporting that PPARγ agonists have the capacity to reduce PH and vascular remodelling in several models of experimental PH, like the monocrotaline-induced or hypoxia-induced mouse models.[35 36] Studies elucidating the mechanisms of PPAR ligand effects in the pulmonary vasculature point to PPAR-mediated alterations in vascular cell proliferation and signalling, progenitor cell function and the production of vasoactive reactive oxygen and nitrogen species.[34] Our findings may have important clinical implications, as at the time of PH diagnosis, the majority of patients have already developed some form of pathological pulmonary arterial remodelling. Therefore, activating PPARs during the pathogenesis of PH or once PH is established holds promise as a therapeutic approach for the disease.

Our study has several limitations that deserve consideration. The preventive setting applied for lung fibrosis in both mouse models may limit the clinical applicability of our results. However, a curative approach was used for PH, since obliteration of pulmonary arteries is usually present at the time IVA337 treatment was initiated.[24] We have also not compared the efficacy of IVA337 to an already used agonist, but similar antifibrotic effects were observed with IVA337 and rosiglitazone in the model of bleomycin-induced dermal fibrosis.[3]

In conclusion, we demonstrate that treatment with 100 mg/kg IVA337 display beneficial effects on inflammatory/immune changes and fibrosis, which are key aspects of SSc. These findings confirm that the pan-PPAR agonist IVA337 is an appealing therapeutic candidate for SSc both for skin and key cardiopulmonary complications.

Acknowledgements Plateformes d'Immuno-biologie(M. Andrieu), de Morphologie/Histologie (M. Favier) et d'imagerie cellulaire (P. Bourdoncle) de l'institutCochin, Paris, France.

Contributors Study design: JA and YA. Conduction of experiments: JA, CG, SP, JS, TG, AC and LT. Data analysis: JA, CG, SP, JS, TG, AC, LT and YA. Writing/drafting and revising the manuscript: JA, IK, CG, SP, JS, TG, AC, LT, J-ML, J-LJ, PB and YA. Final approval of the manuscript: JA, IK, CG, SP, JS, TG, AC, LT, J-ML, J-LJ, PB and YA.

Funding Fundings were supplied by INSERM, ATIP/AVENIR programme and Inventiva. Inventiva was not involved in the study design and data acquisition.

Competing interests JA has/had consultancy relationship and/or has received research funding in relationship with the treatment of systemic sclerosis from Actelion, Roche, Pfizer and Bristol-Myers Squibb. YA has/had consultancy relationship and/or has received research funding in relationship with the treatment of systemic sclerosis from Actelion, Bayer, Biogen Idec, Bristol-Myers Squibb, Genentech/Roche, Inventiva, Medac, Pfizer, Sanofi/Genzyme, Servier and UCB. IK, J-ML and PB are employed by Inventiva. J-LJ has consultancy relationship with Inventiva.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Wynn TA. Cellular and molecular mechanisms of fibrosis. *J Pathol* 2008;214:199–210.
2 Palumbo-Zerr K, Zerr P, Distler A, *et al*. Orphan nuclear receptor NR4A1 regulates transforming growth factor-β signaling and fibrosis. *Nat Med* 2015;21:150–8.
3 Ruzehaji N, Frantz C, Ponsoye M, *et al*. Pan PPAR agonist IVA337 is effective in prevention and treatment of experimental skin fibrosis. *Ann Rheum Dis* 2016;75:2175–83.
4 Ghosh AK, Bhattacharyya S, Wei J, *et al*. Peroxisome proliferator-activated receptor-gamma abrogates Smad-dependent collagen stimulation by targeting the p300 transcriptional coactivator. *Faseb J* 2009;23:2968–77.
5 Samah M, El-Aidy A-R, Tawfik MK, *et al*. Evaluation of the antifibrotic effect of fenofibrate and rosiglitazone on bleomycin-induced pulmonary fibrosis in rats. *Eur J Pharmacol* 2012;689(1-3):186–93.
6 Aoki Y, Maeno T, Aoyagi K, *et al*. Pioglitazone, a peroxisome proliferator-activated receptor gamma ligand, suppresses bleomycin-induced acute lung injury and fibrosis. *Respiration* 2009;77:311–9.
7 Galuppo M, Di Paola R, Mazzon E, *et al*. GW0742, a high affinity PPAR-β/δ agonist reduces lung inflammation induced by bleomycin instillation in mice. *Int J Immunopathol Pharmacol* 2010;23:1033–46.
8 Sutliff RL, Kang BY, Hart CM. PPARgamma as a potential therapeutic target in pulmonary hypertension. *Ther Adv Respir Dis* 2010;4:143–60.
9 Wettstein G, Estivalet C, Tessier J, *et al*. The New Generation Pan-Ppar agonist Iva337 protects the liver from metabolic disorders and fibrosis. *J Hepatol* 2016;64:S169–S170.
10 Herzog EL, Mathur A, Tager AM, *et al*. Review: interstitial lung disease associated with systemic sclerosis and idiopathic pulmonary fibrosis: how similar and distinct? *Arthritis Rheumatol* 2014;66:1967–78.

11 Galiè N, Humbert M, Vachiery JL, *et al*. 2015 ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension: the Joint Task Force for the diagnosis and treatment of pulmonary hypertension of the european Society of Cardiology (ESC) and the european respiratory society (ERS): Endorsed by: association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and lung transplantation (ISHLT). *Eur Respir J* 2015;46:903–75.
12 Humbert M, Sitbon O, Yaïci A, *et al*. Survival in incident and prevalent cohorts of patients with pulmonary arterial hypertension. *Eur Respir J* 2010;36:549–55.
13 Humbert M, Lau EM, Montani D, *et al*. Advances in therapeutic interventions for patients with pulmonary arterial hypertension. *Circulation* 2014;130:2189–208.
14 Avouac J, Elhai M, Tomcik M, *et al*. Critical role of the adhesion receptor DNAX accessory molecule-1 (DNAM-1) in the development of inflammation-driven dermal fibrosis in a mouse model of systemic sclerosis. *Ann Rheum Dis* 2013;72:1089–98.
15 Ponsoye M, Frantz C, Ruzehaji N, *et al*. Treatment with abatacept prevents experimental dermal fibrosis and induces regression of established inflammation-driven fibrosis. *Ann Rheum Dis* 2016;75:2142–9.
16 Elhai M, Avouac J, Hoffmann-Vold AM, *et al*. OX40L blockade protects against inflammation-driven fibrosis. *Proc Natl Acad Sci U S A* 2016;113:E3901–E3910.
17 Guignabert C, Alvira CM, Alastalo TP, *et al*. Tie2-mediated loss of peroxisome proliferator-activated receptor-gamma in mice causes PDGF receptor-beta-dependent pulmonary arterial muscularization. *Am J Physiol Lung Cell Mol Physiol* 2009;297:L1082–L1090.
18 Ricard N, Tu L, Le Hiress M, *et al*. Increased pericyte coverage mediated by endothelial-derived fibroblast growth factor-2 and interleukin-6 is a source of smooth muscle-like cells in pulmonary hypertension. *Circulation* 2014;129:1586–97.
19 Huertas A, Tu L, Thuillet R, *et al*. Leptin signalling system as a target for pulmonary arterial hypertension therapy. *Eur Respir J* 2015;45:1066–80.
20 Le Hiress M, Tu L, Ricard N, *et al*. Proinflammatory signature of the Dysfunctional Endothelium in pulmonary hypertension. role of the macrophage migration inhibitory factor/CD74 complex. *Am J Respir Crit Care Med* 2015;192:983–97.
21 Choi EJ, Jin GY, Bok SM, *et al*. Serial micro-CT assessment of the therapeutic effects of rosiglitazone in a bleomycin-induced lung fibrosis mouse model. *Korean J Radiol* 2014;15:448–55.
22 Genovese T, Cuzzocrea S, Di Paola R, *et al*. Effect of rosiglitazone and 15-deoxy-Delta12,14-prostaglandin J2 on bleomycin-induced lung injury. *Eur Respir J* 2005;25:225–34.
23 Arora MK, Reddy K, Balakumar P. The low dose combination of fenofibrate and rosiglitazone halts the progression of diabetes-induced experimental nephropathy. *Eur J Pharmacol* 2010;636(1-3):137–44.
24 Maurer B, Busch N, Jüngel A, *et al*. Transcription factor fos-related antigen-2 induces progressive peripheral vasculopathy in mice closely resembling human systemic sclerosis. *Circulation* 2009;120:2367–76.
25 Luther J, Ubieta K, Hannemann N, *et al*. Fra-2/AP-1 controls adipocyte differentiation and survival by regulating pparγ and hypoxia. *Cell Death Differ* 2014;21:655–64.
26 Maurer B, Reich N, Juengel A, *et al*. Fra-2 transgenic mice as a novel model of pulmonary hypertension associated with systemic sclerosis. *Ann Rheum Dis* 2012;71:1382–7.
27 Standiford TJ, Keshamouni VG, Reddy RC. Peroxisome proliferator-activated receptor-{gamma} as a regulator of lung inflammation and repair. *Proc Am Thorac Soc* 2005;2:226–31.
28 Ward JE, Gould H, Harris T, *et al*. PPAR gamma ligands, 15-deoxy-delta12,14-prostaglandin J2 and rosiglitazone regulate human cultured airway smooth muscle proliferation through different mechanisms. *Br J Pharmacol* 2004;141:517–25.
29 Keshamouni VG, Reddy RC, Arenberg DA, *et al*. Peroxisome proliferator-activated receptor-gamma activation inhibits tumor progression in non-small-cell lung cancer. *Oncogene* 2004;23:100–8.
30 Lin Q, Fang LP, Zhou WW, *et al*. Rosiglitazone inhibits migration, proliferation, and phenotypic differentiation in cultured human lung fibroblasts. *Exp Lung Res* 2010;36:120–8.
31 Milam JE, Keshamouni VG, Phan SH, *et al*. PPAR-gamma agonists inhibit profibrotic phenotypes in human lung fibroblasts and bleomycin-induced pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2008;294:L891–L901.
32 Takahashi F, Takahashi K, Okazaki T, *et al*. Role of osteopontin in the pathogenesis of bleomycin-induced pulmonary fibrosis. *Am J Respir Cell Mol Biol* 2001;24:264–71.
33 Manoury B, Caulet-Maugendre S, Guénon I, *et al*. TIMP-1 is a key factor of fibrogenic response to bleomycin in mouse lung. *Int J Immunopathol Pharmacol* 2006;19:471–87.
34 Hart CM, Roman J, Reddy R, *et al*. PPARgamma: a novel molecular target in lung disease. *J Investig Med* 2008;56:515–7.
35 Nisbet RE, Sutliff RL, Hart CM. The role of peroxisome proliferator-activated receptors in pulmonary vascular disease. *PPAR Res* 2007;2007:1–10.
36 Lamé MW, Jones AD, Wilson DW, *et al*. Protein targets of monocrotaline pyrrole in pulmonary artery endothelial cells. *J Biol Chem* 2000;275:29091–9.

EXTENDED REPORT

Nintedanib inhibits macrophage activation and ameliorates vascular and fibrotic manifestations in the Fra2 mouse model of systemic sclerosis

Jingang Huang,[1] Christiane Maier,[1] Yun Zhang,[1] Alina Soare,[1] Clara Dees,[1] Christian Beyer,[1] Ulrike Harre,[1] Chih-Wei Chen,[1] Oliver Distler,[2] Georg Schett,[1] Lutz Wollin,[3] Jörg H W Distler[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-2016-210823).

[1]Department of Internal Medicine 3 and Institute for Clinical Immunology, University of Erlangen-Nuremberg, Erlangen, Germany
[2]Rheumaklinik, University Hospital Zurich, Zurich, Switzerland
[3]Boehringer-Ingelheim Pharma, Div. Research Germany, Biberach, Germany

Correspondence to
Dr Jörg H W Distler, Department of Medicine 3 and Institute for Clinical Immunology, University of Erlangen-Nuremberg, Ulmenweg 18, D91054 Erlangen, Germany; joerg.distler@uk-erlangen.de

JH and CM contributed equally.

Received 14 November 2016
Revised 29 May 2017
Accepted 18 July 2017
Published Online First 16 August 2017

ABSTRACT

Background Nintedanib is an inhibitor targeting platelet-derived growth factor receptor, fibroblast growth factor receptor and vascular endothelial growth factor receptor tyrosine kinases that has recently been approved for the treatment of idiopathic pulmonary fibrosis. The aim of this study was to analyse the effects of nintedanib in the fos-related antigen-2 (Fra2) mouse model of systemic sclerosis (SSc).

Methods The effects of nintedanib on pulmonary arterial hypertension with proliferation of pulmonary vascular smooth muscle cells (PVSMCs) and luminal occlusion, on microvascular disease with apoptosis of microvascular endothelial cells (MVECs) and on fibroblast activation with myofibroblast differentiation and accumulation of extracellular matrix were analysed. We also studied the effects of nintedanib on the levels of key mediators involved in the pathogenesis of SSc and on macrophage polarisation.

Results Nintedanib inhibited proliferation of PVSMCs and prevented thickening of the vessel walls and luminal occlusion of pulmonary arteries. Treatment with nintedanib also inhibited apoptosis of MVECs and blunted the capillary rarefaction in Fra2-transgenic mice. These effects were associated with a normalisation of the serum levels of vascular endothelial growth factor in Fra2 mice on treatment with nintedanib. Nintedanib also effectively blocked myofibroblast differentiation and reduced pulmonary, dermal and myocardial fibrosis in Fra2-transgenic mice. The antifibrotic effects of nintedanib were associated with impaired M2 polarisation of monocytes and reduced numbers of M2 macrophages.

Conclusion Nintedanib targets core features of SSc in Fra2-transgenic mice and ameliorates histological features of pulmonary arterial hypertension, destructive microangiopathy and pulmonary and dermal fibrosis. These data might have direct implications for the ongoing phase III clinical trial with nintedanib in SSc-associated interstitial lung disease.

Systemic sclerosis (SSc) is characterised by vascular remodelling with loss of capillary and pulmonary arterial hypertension (PAH) and progressive tissue fibrosis of the skin and internal organs such as the lungs.[1] Although therapies are approved for the treatment of vascular manifestations, targeted therapies for the treatment of fibrosis in SSc are not yet available for clinical use.[2]

Nintedanib, which has recently been approved for the treatment of idiopathic pulmonary fibrosis (IPF), may be an interesting candidate for antifibrotic therapies in SSc. Nintedanib is a potent inhibitor of platelet derived growth factor receptor (PDGFR)-α and PDGFR-β, vascular endothelial growth factor receptor (VEGFR)-1, 2, 3, fibroblast growth factor receptor (FGFR)-1, 2, 3 and SRC family kinases.[3] Recent data demonstrate that nintedanib in pharmacologically relevant concentrations effectively inhibits colony-stimulating factor (CSF)-receptor.[4] All of those pathways have been implicated into the pathogenesis of fibrosis and have been discussed as candidates for targeted therapies in SSc.[5–8] Nintedanib may thus offer the potential to simultaneously inhibit multiple profibrotic pathways with a single drug. Indeed, we demonstrated potent antifibrotic effects of nintedanib in several preclinical models of skin fibrosis[9] and a phase III clinical trial with nintedanib in SSc-associated interstitial lung disease is currently ongoing (SENSCIS trial; NCT02597933). However, the effects of nintedanib on vascular manifestations have not been investigated so far.

While its inhibitory effects on PDGFRs may be beneficial in PAH, the effects of nintedanib on microvascular disease in SSc are of concern. Given that PDGF, VEGF and FGF signalling as well as SRC kinases have all been implicated in angiogenesis,[10 11] nintedanib may interfere with vascular repair and regeneration. Although no adverse events related to inhibition of angiogenesis have been observed in clinical trials with patients with IPF,[12] patients with SSc might be more sensitive due to the pre-existing microvascular disease. On the other hand, accumulating evidence suggests that uncontrolled VEGF signalling in SSc may actually promote microvascular disease and capillary loss in SSc.[5 7]

Fos-related antigen-2 transgenic (Fra2 transgenic) mice resemble the core clinical features of SSc as they display fibrotic and vascular manifestations of SSc.[13–16] Fra2 transgenic mice develop a destructive microvascular disease with apoptosis of endothelial cells followed by systemic fibrotic manifestations and PAH. In this study, we used Fra2 transgenic mice as a model system to study the effects of nintedanib on vascular and fibrotic manifestations of SSc.

To cite: Huang J, Maier C, Zhang Y, *et al*. *Ann Rheum Dis* 2017;**76**:1941–1948.

MATERIALS AND METHODS

Fra2 transgenic mice

Treatment of Fra2 transgenic mice with nintedanib was initiated at an age of 10 weeks and mice were sacrificed at an age of 16 weeks. Non-transgenic littermates served as controls. Each group consisted of six mice. Nintedanib was administered by oral gavage. Controls received vehicle treatment.

Quantification of pulmonary, dermal and cardiac fibrosis

Skin fibrosis was assessed histologically and in addition by quantification of the hydroxyproline content and analyses of myofibroblast counts. Interstitial lung changes were analysed histologically by H&E and trichrome staining, by quantification of myofibroblasts, by immunofluorescence and biochemically by determination of hydroxyproline contents.[15 17–20] Assessment of cardiac fibrosis also followed established protocols.[21]

Evaluation of vascular remodelling of the pulmonary arteries

Vascular changes were evaluated according to the Dana Point consensus criteria using established protocols and readouts.[22 23] Muscular blood vessels were identified by positive staining for Smooth Muscle Protein 22-alpha, green (α-SMA) and α smooth muscle actin (SM22α). While α-SMA is also expressed by myofibroblasts, the expression of SM22α is restricted to vascular smooth muscle cells. The vessel wall thickness of pulmonary arteries was measured on H&E stained sections captured at 400-fold magnification. For analysis, only the circular-shaped vessels were included, the oblique and longitudinal sectioned vessels were excluded. The thickness of vessel walls was evaluated by a minimum of three measurements per vessel[14] by using the following equation: (outside diameter-inside diameter)/outside diameter. The results were express as x-fold changes.[24] As an additional outcome for vascular remodelling, the degree of luminal occlusion of pulmonary arteries (α-SMA and SM22α-positive) was examined by counting the numbers of occluded lumina/high-power field (HPF).[25] The percentage of proliferating vascular smooth muscle cells was evaluated by triple staining for 4′,6-Diamidin-2-phenylindol (DAPI) (visualisation of nuclei), Ki67 (proliferation marker) and SM22α.[13 14 16]

Assessment of microvascular changes

The percentage of apoptotic endothelial cells in the skin of Fra2 mice was determined by triple staining with the endothelial cell (EC) marker CD31, DAPI and with the TdT-mediated dUTP-biotin nick end labeling (TUNEL) assay[14 26] in five HPF per mouse.[14 27]

IMMUNOFLUORESCENCE STAININGS

M2 macrophages can be identified by immunohistochemistry using a combination of pan-macrophage markers (F4/80 in mice) and prototypical M2 markers such as arginase and cMAF, a transcription factor that has recently been shown to be required for the expression of arginase.[28] M1 macrophages were defined by staining for inducibile nitric oxide synthase (iNOS), CD11c and F4/80. According to established protocols,[29] the following primary antibodies were used: F4/80 (Biorad, Puchheim, Germany), arginase (Santa Cruz, San Diego, California, USA), cMAF (Abgent, San Diego, California, USA) iNOS (Invitrogen) and CD11c (Abcam, Cambridge, UK). In a subset of experiments, costainings of F4/80 with either interleukin (IL)-12 (LSBIO, Seattle, Washington, USA), IL-4 (Santa Cruz) or IL-13 (Bioss, Woburn, Massachusetts, USA) were performed. Additional stainings included vimentin (mesenchymal cells, Merck, Darmstadt, Germany) and CD45 (leukocytes (Abcam)). The following secondary antibodies were used: Alexa Fluor 647 chicken anti-rat (Invitrogen, Germany), Alexa Fluor 488 donkey anti-rabbit (Abcam), Alexa Fluor 594 donkey anti-goat (Invitrogen).[20] Sections were counterstained with DAPI.

Quantification of the staining intensity

The staining intensity was quantified as described using ImageJ.[14] A density threshold was set to quantify the positive staining by using the respective negative controls. The threshold was selected to exclude unspecific background staining. The same thresholds and system settings were used for all slides. The number of pixels falling within the threshold, indicating the quantity of staining positivity, was recorded for each field.

Isolation, selection and culture of macrophages

Peripheral blood mononuclear cells (PBMCs) were isolated from six healthy volunteers using Lymphoflot (Bio-Rad, Hercules, California, USA). CD14-microbeads (Miltenyi Biotech, Bergisch-Gladbach, Germany) were used for positive selection of human monocytes and macrophages.[26] Monocytes/macrophages were seeded in RPMI medium supplemented with 10% fetal bovine serum (FBS) (all Gibco, Basel, Switzerland) with or without 25 nM human M-CSF (Peprotech, Rocky Hill, New Jersey, USA) for 6 days at 37°C and 5% CO_2. At days 3 and 6, the medium was replaced by fresh medium. At day 6, cells were treated with 100 nM of nintedanib. Twenty-four hours later, cells were stimulated with 10 nM IL−4 and 10 nM IL-13 (both Peprotech). Twenty-four hours after stimulation, the expression of M1 and M2 markers was measured.

Microtitre tetrazolium assay

The microtitre tetrazolium [3, (4, 5-dimethylthiazol-2-yl)2, 5-diphenyl-tetrazolium-bromide] assay is an established method to analyse proliferation. The effect of nintedanib on primary human vascular smooth muscle cells was analysed in the presence or absence of PDGF.

Cytokine measurements

Cytokines in the serum of mice were quantified using the mouse inflammation multianalyte profile, which is based on ultrasensitive immunoassays for 37 cytokines (Myriad-RBM, Austin, Texas, USA).

Statistics

All data are presented as median±IQR. Differences were analysed using the Mann-Whitney U test. p Values are expressed as follows: $0.05 > p \geq 0.01$ as *; $0.01 > p \geq 0.001$ as **; $p < 0.001$ as ***.

RESULTS

Treatment with nintedanib reduces dermal, pulmonary and myocardial fibrosis in Fra2 transgenic mice

We first evaluated the effect of nintedanib on fibrosis of the skin and the lungs. Treatment with nintedanib was well tolerated and reduced the weight loss in Fra2 transgenic mice (see online supplementary figure S1). We did not observe changes in activity, in texture of the fur or in the consistency of the stool at both doses.

Doses of 60 mg/kg qd or 50 mg/kg/two times per day strongly ameliorated skin fibrosis. Both doses of nintedanib were equally effective and significantly reduced dermal thickening, collagen

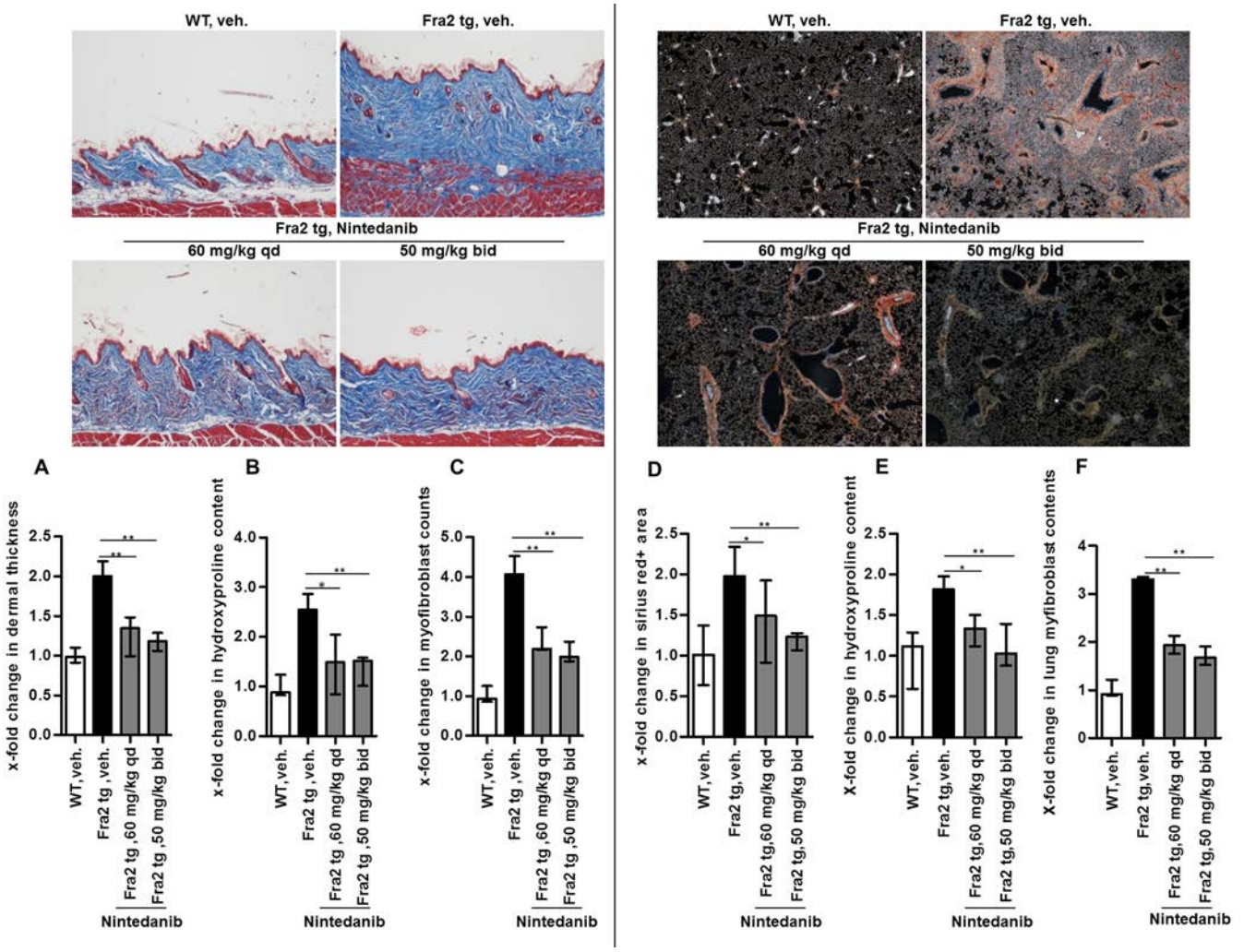


Figure 1 Nintedanib reduces dermal and pulmonary fibrosis in fos-related antigen-2 (Fra2) transgenic mice. (A–C) Effects on skin fibrosis. Representative images of Masson trichrome-stained skin sections at 100-fold magnification and quantification of dermal thickness (A), hydroxyproline content (B) and myofibroblast count (C). (D–F) Effects on pulmonary fibrosis. Representative images of Sirius red-stained sections of lungs at 40-fold magnification and quantification of the fibrotic area (D), of the hydroxyproline content (E) and of myofibroblast counts (F). *0.01≤p<0.05; **0.001≤p<0.01 versus vehicle-treated Fra2-transgenic mice (Fra2 tg). WT, vehicle-treated wild-type mice, n=6 mice per group.

accumulation and myofibroblast differentiation as compared with vehicle-treated Fra2 transgenic mice (figure 1A–C).

Nintedanib-treated mice also demonstrated reduced pulmonary fibrosis with decreased fibrotic area, lower hydroxyproline content and decreased myofibroblast counts compared with control mice (figure 1D–F).

Myocardial changes in patients with SSc are mimicked in Fra2 transgenic mice.[21] Treatment with nintedanib at 60 mg/kg/qd reduced the extent of fibrosis, ameliorated perivascular inflammation and decreased apoptosis of endothelial cells (see online supplementary figure S2A–F).

Nintedanib inhibits remodelling of the pulmonary arteries

Fra2 transgenic mice also develop pulmonary arterial hypertension with extensive remodelling of the pulmonary arteries. Both doses of nintedanib significantly reduced thickening of the walls of pulmonary arteries (figure 2A, B) and decreased the number of occluded vessels compared with vehicle-treated Fra2 transgenic mice (figure 2C). Consistent with these findings, the number of proliferating vascular smooth muscle cells was strongly decreased in nintedanib-treated mice (figure 2D, E). We observed a trend towards higher efficacy with nintedanib at 50 mg/kg/two times per day as compared with 60 mg/kg/qd, which, however, did not reach statistical significance. Nintedanib in a concentration-dependent manner also inhibited proliferation of human pulmonary vascular smooth muscle cells in vitro, both under basal conditions and on stimulation with PDGF (figure 2F).

Nintedanib ameliorates microvascular disease in Fra2 transgenic mice

Another characteristic feature of Fra2 transgenic mice is apoptosis of endothelial cells with subsequent loss of capillaries. The number of apoptotic endothelial cells in the skin and lungs of Fra2 transgenic mice was reduced by treatment with nintedanib (figure 3A, B and online supplementary figure S3). Consistently, capillary loss was reduced and increased numbers of vessels in the skin were observed in the dermis of nintedanib-treated mice (figure 3C). As for remodelling of pulmonary arteries, doses of 50 mg/kg/two times per day tended to be more effective than 60 mg/kg/qd and statistically significant

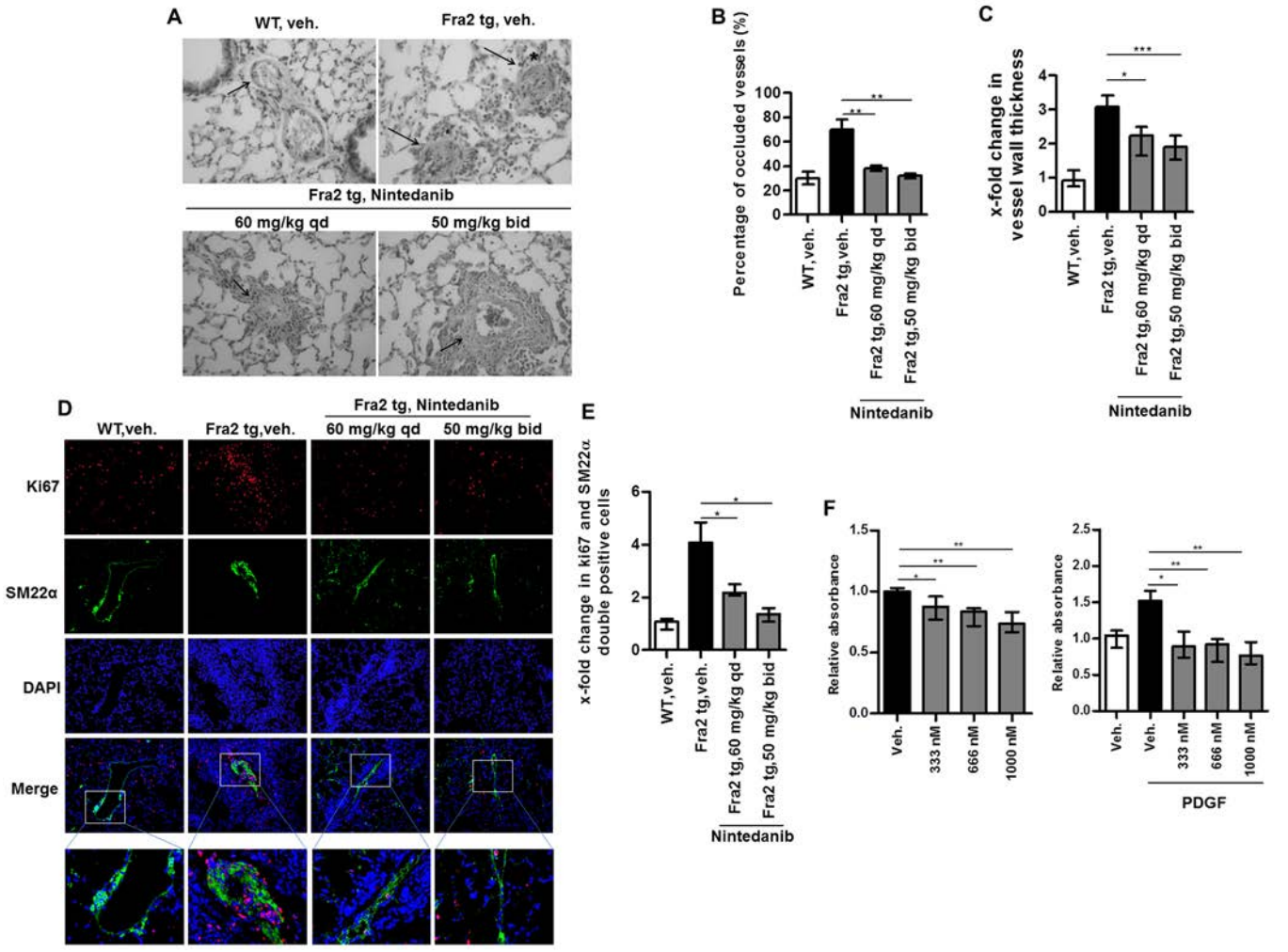


Figure 2 Treatment with nintedanib reduces vascular remodelling of the pulmonary arteries in Fra2 transgenic mice. (A) Representative images of H&E-stained lung sections with thickened vessel walls shown at 400-fold magnification. Arrows indicate thickened vessel walls, stars indicate occluded vessels. (B) Percentage of occluded vessels. (C) Average thickness of the vessel wall. (D) Representative examples of triple staining for DAPI (nuclear staining), Ki67 (proliferation marker) and SM22α (vascular smooth muscle cells) at 200-fold and 600-fold magnification. (E) Relative proportion of proliferating vascular smooth muscle cells per total number of vascular smooth muscle cells. (F) Effects of nintedanib on the proliferation of human pulmonary artery vascular smooth muscle cells at basal conditions and on stimulation with platelet-derived growth factor (PDGF). *0.01≤p<0.05; **0.001≤p<0.01; ***p<0.001 versus vehicle-treated Fra2 transgenic mice. DAPI, 4′,6-diamidin-2-phenylindol, blue; Fra2 tg, Fra2-transgenic mice; SM22α, smooth muscle protein 22-alpha, green; WT, vehicle-treated wild-type mice.

differences were found for total microvessel counts between both groups.

Treatment with Nintedanib normalises the levels of M-CSF1 and VEGF in Fra2 transgenic mice

To gain additional insights into the mechanisms underlying the anti-fibrotic and vasoprotective effects of nintedanib in the Fra2 model, we screened for differences in the serum levels of central inflammatory, fibrotic and angiogenic mediators. Several of these cytokines and growth factors such as epithelial growth factor (EGF), IL-1β, IL-5, IL-6, IL-10, IL-18, IP-10, MCP-1, MCP-3, MCP-5, M-CSF, macrophage inflammatory protein (MIP)-1α, MIP-1β, stem cell factor (SCF), tissue inhibitor of MMP (TIMP)-1, thrombopoietin and VEGF were upregulated in Fra2 transgenic mice as compared with non-transgenic littermates (figure 4A). Of those, two mediators were significantly affected by treatment with nintedanib. Treatment with nintedanib decreased the levels of M-CSF (figure 4B) and of VEGF (at a dose of 60 mg/kg/qd) in Fra2 transgenic mice (figure 4C).

Nintedanib inhibits M2 differentiation of monocytes

M-CSF is a central growth factor for monocytes and macrophages and favours their alternative activation and differentiation into M2 macrophages.[30] M2 macrophages are implicated into the pathogenesis of fibrosis.[31 32] Changes in the expression of M2 genes have recently been linked to clinical responses in tocilizumab-treated patients with SSc.[33 34] Based on the down-regulation of M-CSF, we hypothesised that nintedanib may inhibit M2 polarisation in Fra2 transgenic mice. Indeed, we observed highly increased numbers of M2 macrophages in the skin of Fra2 transgenic mice compared with wildtype (WT) littermates. Treatment with nintedanib completely prevented the increase in M2 macrophages with M2 counts comparable to those in non-transgenic control (figure 4D and online supplementary figure S4). In contrast to M2 macrophages, the number of M1 macrophages (defined by the expression of iNOS and CD11c in combination with F4/80) did not differ between Fra2

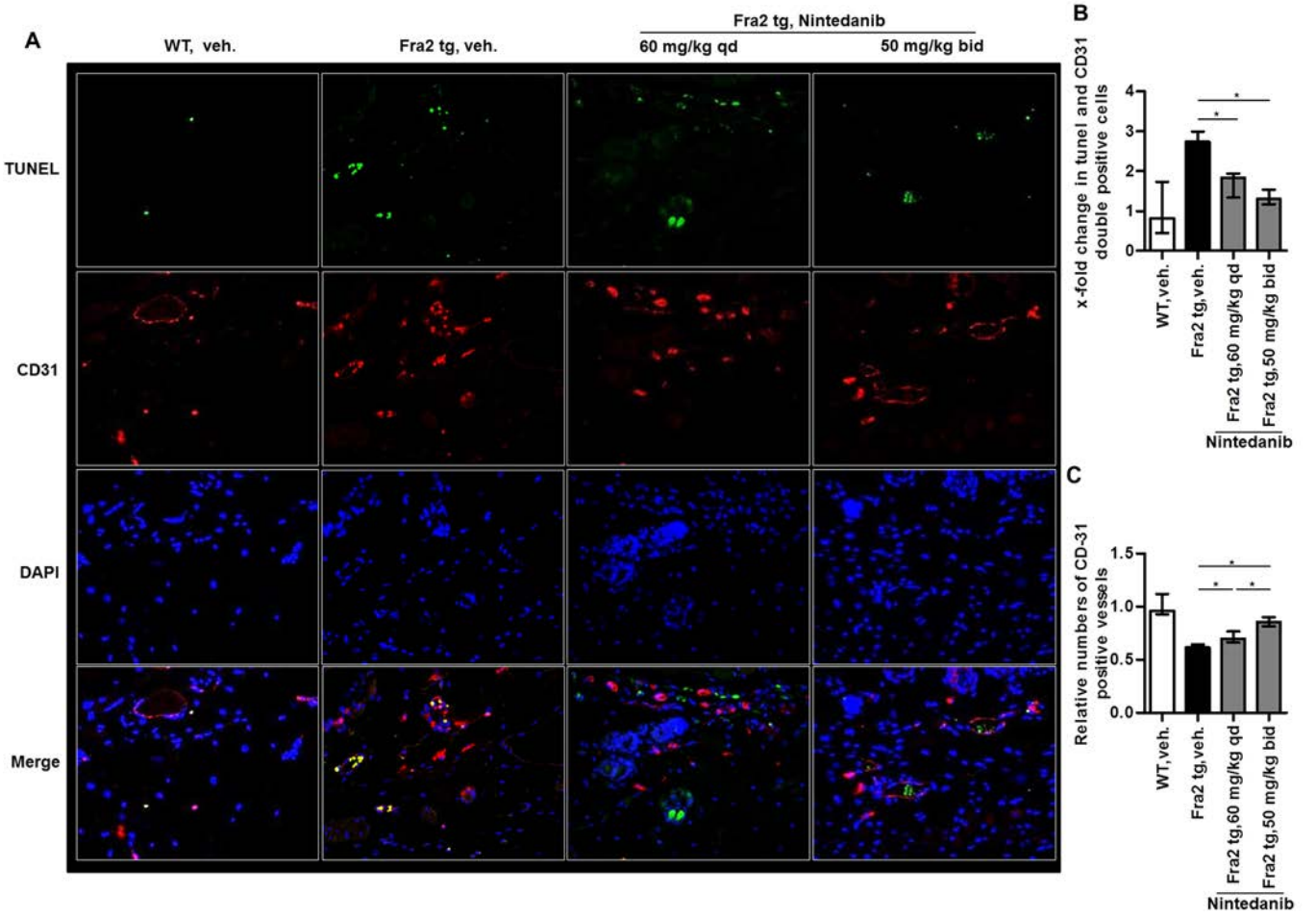


Figure 3 Nintedanib ameliorates microvascular manifestations in fos-related antigen-2 (Fra2) transgenic mice. (A) Representative images of skin sections stained for TdT-mediated dUTP-biotin nick end labeling (TUNEL) (apoptosis marker), CD31 (endothelial cells) and 4′,6-diamidin-2-phenylindol (DAPI) (nuclear staining) shown at 400-fold magnification. (B) Quantification of apoptotic microvascular endothelial cells in the skin. (C) Relative numbers of CD31-positive vessels. n=6 mice per group. *0.01≤p<0.05 versus vehicle-treated Fra2 transgenic mice. Fra2 tg, Fra2-transgenic mice; WT, vehicle-treated wild-type mice.

transgenic mice and non-transgenic littermates and M1 macrophage counts were not affected by treatment with nintedanib (figure 4E).

To confirm that inhibitory effects of nintedanib are not restricted to the Fra2 model or to murine monocytes, we tested the effects of nintedanib on alternative activation of human monocytes induced by M-CSF1, IL-4 and IL-13. Nintedanib reduced M2 counts with suppression of the expression levels of individual M2 markers such as CD163 or CD206 (figure 5A). The expression of M1 markers such as CD86, toll-like receptor 4 (TLR4) and human leukocyte antigen-DR (HLA-DR) (figure 5B) or general markers of monocytes such as CD14 was unchanged or even increased by incubation with nintedanib (data not shown). The effects of nintedanib on macrophage polarisation could be explained by its direct inhibitory effects on CSF1R (recent, unpublished information by Boehringer Ingelheim). In support of this hypothesis, individual inhibition of PDGFR, FGFR or VEGFR did not affect M2 polarisation of macrophages (see online supplementary file S5A and B). Together, these data demonstrate that nintedanib inhibits M2 polarisation of macrophages, likely by direct inhibition of CSF1R.

To confirm the functional relevance of those findings, we analysed the release of M1 and M2 cytokines in Fra2 transgenic mice. The numbers of macrophages positive for IL-4 or for IL-13 were strongly increased in Fra2 transgenic mice. Treatment with nintedanib reduced the number of IL-4 and IL13-positive macrophages (figure 5C) back to the levels of non-transgenic mice. In contrast, the number of IL-12 positive macrophages was not increased in Fra2-transgenic mice and was not affected by treatment with nintedanib (figure 5C). Similar results were obtained for the total numbers of IL-4, IL-13 and IL-12 positive cells (data not shown).

DISCUSSION

Fra2 transgenic mice resemble the clinical manifestations of SSc and develop not only systemic fibrotic disease, but also SSc-like microvascular disease and PAH.[13–16] Evidence provided by studies on the efficacy of imatinib, which rather selectively targets PDGF signalling and cellular abelson kinase (c-ABL) in fibrotic diseases, indicates that Fra2 transgenic mice may reflect the activation levels of profibrotic pathways in human SSc more closely and may better predict responses as compared with other mouse models.[23 35–37] Indeed, the tyrosine kinase inhibitor imatinib that did not reach its primary end point in clinical trials in SSc was also not effective in Fra2 transgenic mice, but reduced fibrosis induced by bleomycin.[23] We now demonstrate that nintedanib, which inhibits FGFRs, PDGFRs, VEGFRs, SRC family kinases and, according to very recent data, also CSF1R, effectively reduces dermal and pulmonary fibrosis in Fra2 transgenic mice,

A

Cytokine	WT mice	Fra2 mice	P-value
EGF	252 ± 10	315 ± 8	0.0217
IL-1β	9513 ± 478	13800 ± 372	0.0010
IL-5	840 ± 80	1780 ± 159	0.0246
IL-6	not detectable	24 ± 6	-
IL-10	not detectable	236 ± 14	-
IL-18	25125 ± 1827	33600 ± 1079	0.0012
IP-10	74 ± 5	300 ± 27	0.0113
MCP-1	73 ± 9	205 ± 15	0.0064
MCP-3	212 ± 37	674 ± 44	0.0032
MCP-5	27 ± 5	44 ± 4	0.0019
MIP-1α	7038 ± 369	9040 ± 205	0.0399
MIP-1β	182 ± 11	445 ± 35	0.0120
SCF	762 ± 34	1227 ± 78	0.0206
TIMP-1	1264 ± 136	9700 ± 1221	0.0364
Thrombopoietin	56750 ± 2202	86400 ± 2115	< 0.0001

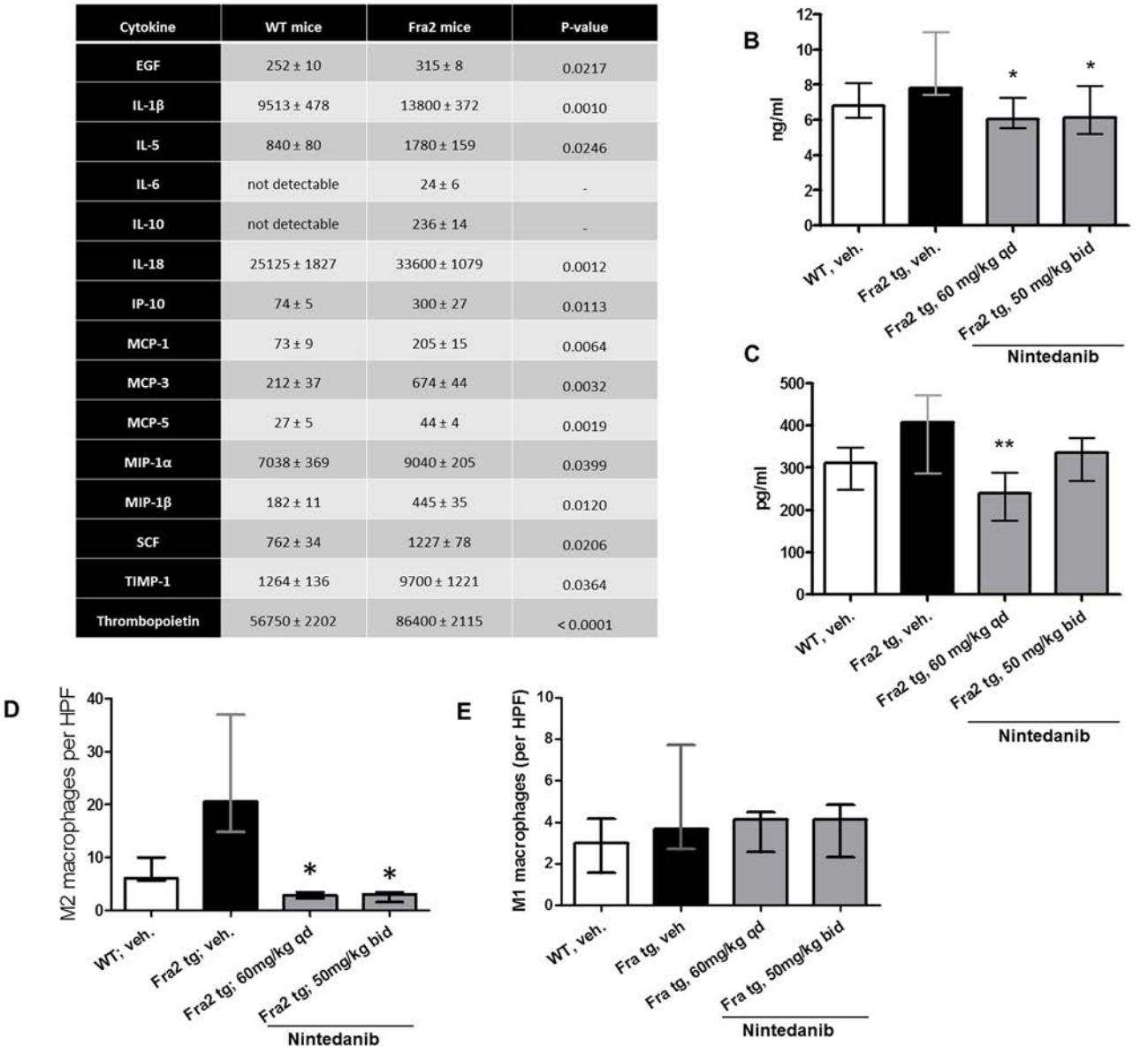


Figure 4 Treatment with nintedanib alters the levels of macrophage (M)-colony-stimulating factor (CSF) and vascular endothelial growth factor (VEGF) and reduces M2 macrophage counts in fos-related antigen-2 (Fra2) transgenic mice. (A) Serum levels of major inflammatory, fibrotic and angiogenic mediators in Fra2 transgenic mice and littermate controls in pg/mL. Effects of nintedanib on the levels of M-CSF1 (B) and VEGF (C) in Fra2 transgenic mice. n=6 mice per group. (D) Fold changes in M2 macrophage counts defined as F4/80, cMAF, arginase triple-positive cells. (E) Fold changes in M1 macrophage counts defined as F4/80, iNOS, CD11c triple-positive cells. *0.01≤p<0.05; **0.001≤p<0.01 versus vehicle-treated Fra2 transgenic mice. Fra2 tg, Fra2-transgenic mice; HPF, high-power field; IL, interleukin; WT, vehicle-treated wild-type mice.

thereby extending previous results on cultured fibroblasts and localised skin fibrosis.[9]

In addition to its direct effects on fibroblasts, we provide here novel evidence that nintedanib may also target fibroblast activation indirectly by blocking alternative activation and M2 polarisation of macrophages. Nintedanib inhibits M2 polarisation of healthy human macrophages in vitro and also strongly reduces M2 macrophage counts in Fra2 transgenic mice. These inhibitory effects of nintedanib on macrophage polarisation are likely mediated by direct inhibition of CSF1R. Selective inhibition of PDGFR, VEGFR or FGFR did not block M2 polarisation in vitro, whereas targeted inhibition of CSF1R is known to interfere with the alternative activation of macrophages.[38] However, confirmation of the findings with monocytes from patients with SSc is warranted. Given that M2 macrophages are a rich source of profibrotic mediators[34] and M2 macrophages have been shown to strongly affect the outcome in various preclinical models of fibrosis[39] and that changes in M2 mRNA signature have recently been linked to clinical outcomes in patients with SSc,[20 40] the effects of nintedanib on M2 polarisation might significantly contribute to the antifibrotic effects of nintedanib in vivo and may be of direct clinical relevance.

Although our study is limited by the lack of functional assessment by echocardiography or right heart catheter, we provide clear evidence that treatment with nintedanib also inhibited the proliferation of pulmonary vascular smooth muscle cells and ameliorated the histological features of PAH. Further studies are required to unravel whether those effects are solely mediated by inhibition of PDGFR or whether other targets of nintedanib contribute to the potent inhibitory effects. Despite the technical challenges of right heart catheterisation in mice, i.p. with widespread organ involvement as in Fra2 transgenic mice, right heart

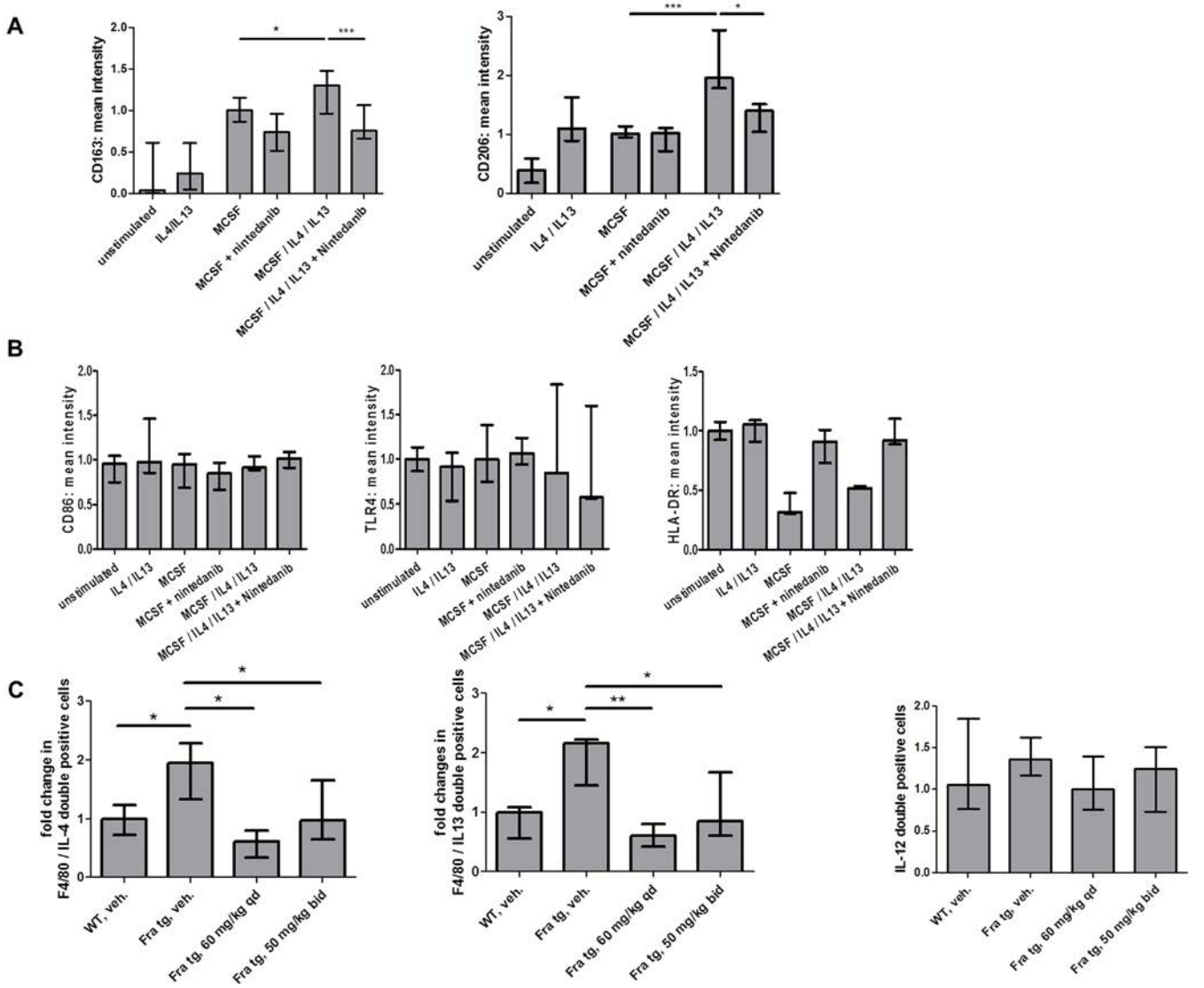


Figure 5 Nintedanib inhibits alternative activation of monocytes. (A) Effects of nintedanib on M2 polarisation of human macrophages incubated with macrophage (M)-colony-stimulating factor (CSF), interleukin (IL)-4 and IL-13 as analysed by changes in the mean fluorescence intensity of CD163and CD206. (B) Expression levels of M1 markers CD86,toll-like receptor 4 (TLR4) and human leukocyte antigen-DR (HLA-DR). (C) Fold changes in the number of IL-4, IL-13 and IL-12 positive macrophages in fos-related antigen-2 (Fra2) transgenic mice with or without nintedanib treatment. *$0.01 \leq p < 0.05$; **$0.001 \leq p < 0.01$.

catheterisation in follow-up studies would further confirm the beneficial effects of nintedanib on PAH. At the molecular level, inhibition of PDGFR with impaired proliferation of vascular smooth muscle cells is likely to play a central role for the beneficial effects of nintedanib on PAH. As recent data provide elegant evidence that apoptosis of endothelial cell may promote PAH,[39] further studies should investigate whether the positive effects of nintedanib on the histological features of PAH are in part mediated by inhibition of pulmonary vascular EC apoptosis.

Treatment with nintedanib also reduced apoptosis of microvascular endothelial cells of the skin and ameliorated the loss of capillaries in Fra2 transgenic mice. These findings seem unexpected on first view given the inhibition of multiple proangiogenic mediators by nintedanib. However, excessive upregulation of angiogenic factors such as VEGF has previously been shown to perturb, rather than to promote angiogenesis and to induce microvascular alterations reminiscent of those observed in SSc.[5 7] Treatment with nintedanib may thus improve microvascular manifestations in Fra2 transgenic mice by preventing the deleterious effects of excessive, uncontrolled activity of angiogenic factors. This hypothesis is supported by the finding that the levels of VEGF are upregulated in Fra2 transgenic mice and that treatment with nintedanib in doses of 60 mg/kg/qd normalised the levels of VEGF. Mechanistically, nintedanib may inhibit VEGF expression by inhibiting inflammation and in particular macrophage activation, as macrophages are a major source of VEGF.[41 42] The effects of nintedanib on M-CSF may have also contributed to the beneficial effects on vascular alterations. M-CSF has been shown to be capable of modulating angiogenesis directly, but can also regulate the activation of endothelial cells indirectly by inducing the expression of VEGF.[43] However, further studies are required to decipher the molecular regulation of VEGF by nintedanib. Moreover, it will be crucial to assess the effects of nintedanib on angiogenesis during wound healing that demands increased formation of new vessels to initiate tissue repair.

Apart from potential effects on physiological wound healing, common side effects of nintedanib with particular relevance to SSc are gastrointestinal adverse events. Although mild in most patients,

gastrointestinal side effects were common in IPF studies and may complicate gastrointestinal involvement in SSc.

In summary, our data provide preclinical evidence that treatment with nintedanib may not only improve fibrosis of skin, lungs and heart, but also ameliorate vascular manifestations as the other major cause of morbidity and mortality in SSc. We also provide evidence that nintedanib may not only exert its antifibrotic effects by direct inhibition of fibroblast activation, but also by inhibition of alternative activation of macrophages. It will be important to follow-up on these preclinical findings in the ongoing clinical study and to carefully monitor microvascular disease and PAH.

Correction notice This article has been corrected since it published Online First. The title has been corrected to: Nintedanib inhibits macrophage activation and ameliorates vascular and fibrotic manifestations in the Fra2 mouse model of systemic sclerosis.

Acknowledgements We thank Regina Kleinlein and Katja Dreißigacker for excellent technical assistance.

Contributors Design of the study: LW, JHWD. Acquisition of data: JH, CM, YZ, AS, CD, CB, CWC. Interpretation of data: JH, CM, YZ, AS, CD, CB, CWC, UH, OD, GS, LW, JHWD. Manuscript preparation: JH, CM, LW, JHWD.

Funding Grants DI 1537/5-1, DI 1537/7-1, DI 1537/8-1, DI 1537/9-1, DI 1537/11-1, DE 2414/2-1 andAK 144/2-1 of the Deutsche Forschungsgemeinschaft, grants A57, and A64 of the IZKF inErlangen, grant 2013.056.1 of the Wilhelm-Sander-Foundation, grants 2014_A47 and2014_A248 of the Else-Kröner-Fresenius-Foundation, grant 2013.056.1 of the Wilhelm SanderFoundation and a Career Support Award of Medicine of the Ernst Jung Foundation.

Competing interests OD has consultancy relationships and/or has received research funding from Actelion, Pfizer, Ergonex, BMS, Sanofi-Aventis, United BioSource Corporation, Roche/Genentech, Medac, Biovitrium, Boehringer Ingelheim, Novartis, 4D Science, Active Biotech, Bayer, Sinoxa, Serodapharm, EpiPharm, GSK, Pharmacyclics and Biogen. LW is an employee of Boehringer-Ingelheim. JHWD has consultancy relationships with Actelion, Active Biotech, Anamar, Bayer Pharma, Boehringer Ingelheim, Celgene, Galapagos, GSK, Inventiva, JB Therapeutics, Medac, Pfizer, RuiYi and UCB. JHWD has received research funding from Anamar, Active Biotech, Array Biopharma, BMS, Bayer Pharma, Boehringer Ingelheim, Celgene, GSK, Novartis, Sanofi-Aventis and UCB. JHWD is stock owner of 4D Science.

Patient consent Obtained.

Ethics approval Ethical Committee of the University of Erlangen-Nuremberg.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. *N Engl J Med* 2009;360:1989–2003.
2 Ramming A, Dees C, Distler JH. From pathogenesis to therapy--Perspective on treatment strategies in fibrotic diseases. *Pharmacol Res* 2015;100:93–100.
3 Hilberg F, Roth GJ, Krssak M, *et al*. BIBF 1120: triple angiokinase inhibitor with sustained receptor blockade and good antitumor efficacy. *Cancer Res* 2008;68:4774–82.
4 Tandon F, Herrmann E, Ayaub P, *et al*. Nintedanib attenuates the polarization of profibrotic macrophages through the inhibition of tyrosine phosphorylation on CSF1 receptor K. *Am J Respir Crit Care Med* 2017;195:A2397.
5 Distler O, Distler JH, Scheid A, *et al*. Uncontrolled expression of vascular endothelial growth factor and its receptors leads to insufficient skin angiogenesis in patients with systemic sclerosis. *Circ Res* 2004;95:109–16.
6 Olson LE, Soriano P. Increased PDGFRalpha activation disrupts connective tissue development and drives systemic fibrosis. *Dev Cell* 2009;16:303–13.
7 Maurer B, Distler A, Suliman YA, *et al*. Vascular endothelial growth factor aggravates fibrosis and vasculopathy in experimental models of systemic sclerosis. *Ann Rheum Dis* 2014;73:1880–7.
8 Skhirtladze C, Distler O, Dees C, *et al*. Src kinases in systemic sclerosis: central roles in fibroblast activation and in skin fibrosis. *Arthritis Rheum* 2008;58:1475–84.
9 Huang J, Beyer C, Palumbo-Zerr K, *et al*. Nintedanib inhibits fibroblast activation and ameliorates fibrosis in preclinical models of systemic sclerosis. *Ann Rheum Dis* 2016;75:150.3–1.
10 Yoo SY, Kwon SM. Angiogenesis and its therapeutic opportunities. *Mediators Inflamm* 2013;2013:1–11.
11 Distler JH, Hirth A, Kurowska-Stolarska M, *et al*. Angiogenic and angiostatic factors in the molecular control of angiogenesis. *Q J Nucl Med* 2003;47:149–61.
12 Richeldi L, du Bois RM, Raghu G, *et al*. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2071–82.
13 Maurer B, Busch N, Jungel A, *et al*. Transcription factor fos-related antigen-2 induces progressive peripheral vasculopathy in mice closely resembling human systemic sclerosis the transcription factor Fra-2 regulates the production of extracellular matrix in systemic sclerosis. *Circulation* 2009;120:2367–76.
14 Maurer B, Reich N, Juengel A, *et al*. Fra-2 transgenic mice as a novel model of pulmonary hypertension associated with systemic sclerosis. *Ann Rheum Dis* 2012;71:1382–7.
15 Reich N, Maurer B, Akhmetshina A, *et al*. The transcription factor Fra-2 regulates the production of extracellular matrix in systemic sclerosis. *Arthritis Rheum* 2010;62:280–90.
16 Eferl R, Hasselblatt P, Rath M, *et al*. Development of pulmonary fibrosis through a pathway involving the transcription factor Fra-2/AP-1. *Proc Natl Acad Sci U S A* 2008;105:10525–30.
17 Akhmetshina A, Palumbo K, Dees C, *et al*. Activation of canonical wnt signalling is required for TGF-β-mediated fibrosis. *Nat Commun* 2012;3:735.
18 Dees C, Akhmetshina A, Zerr P, *et al*. Platelet-derived serotonin links vascular disease and tissue fibrosis. *J Exp Med* 2011;208:961–72.
19 Palumbo-Zerr K, Zerr P, Distler A, *et al*. Orphan nuclear receptor NR4A1 regulates transforming growth factor-β signaling and fibrosis. *Nat Med* 2015;21:150–8.
20 Avouac J, Palumbo K, Tomcik M, *et al*. Inhibition of activator protein 1 signaling abrogates transforming growth factor β-mediated activation of fibroblasts and prevents experimental fibrosis. *Arthritis Rheum* 2012;64:1642–52.
21 Venalis P, Kumánovics G, Schulze-Koops H, *et al*. Cardiomyopathy in murine models of systemic sclerosis. *Arthritis Rheumatol* 2015;67:508–16.
22 Pietra GG, Capron F, Stewart S, *et al*. Pathologic assessment of vasculopathies in pulmonary hypertension. *J Am Coll Cardiol* 2004;43(12 Suppl S):S25–S32.
23 Maurer B, Distler A, Dees C, *et al*. Levels of target activation predict antifibrotic responses to tyrosine kinase inhibitors. *Ann Rheum Dis* 2013;72:2039–46.
24 Dorfmüller P, Humbert M, Capron F. [Update on the pathomorphological assessment of vasculopathies in pulmonary arterial hypertension]. *Pathologe* 2006;27:140–6.
25 al-Sabbagh MR, Steen VD, Zee BC, *et al*. Pulmonary arterial histology and morphometry in systemic sclerosis: a case-control autopsy study. *J Rheumatol* 1989;16:1038–42.
26 Palumbo-Zerr K, Zerr P, Distler A, *et al*. Orphan nuclear receptor NR4A1 regulates transforming growth factor-β signaling and fibrosis. *Nat Med* 2015;21:150–8.
27 Distler JH, Jüngel A, Pileckyte M, *et al*. Hypoxia-induced increase in the production of extracellular matrix proteins in systemic sclerosis. *Arthritis Rheum* 2007;56:4203–15.
28 Nakamura M, Hamada M, Hasegawa K, *et al*. c-Maf is essential for the F4/80 expression in macrophages in vivo. *Gene* 2009;445:66–72.
29 Barros MH, Hauck F, Dreyer JH, *et al*. Macrophage polarisation: an immunohistochemical approach for identifying M1 and M2 macrophages. *PLoS One* 2013;8:e80908.
30 Schett G. Review: immune cells and mediators of inflammatory arthritis. *Autoimmunity* 2008;41:224–9.
31 Wynn TA, Barron L. Macrophages: master regulators of inflammation and fibrosis. *Semin Liver Dis* 2010;30:245–57.
32 Wynn TA, Vannella KM. Macrophages in tissue repair, regeneration, and fibrosis. *Immunity* 2016;44:450–62.
33 Distler O, Distler JH. Tocilizumab for systemic sclerosis: implications for future trials. *Lancet* 2016;387:2580–1.
34 Khanna D, Denton CP, Jahreis A, *et al*. Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial. *Lancet* 2016;387:2630–40.
35 Akhmetshina A, Dees C, Pileckyte M, *et al*. Dual inhibition of c-abl and PDGF receptor signaling by dasatinib and nilotinib for the treatment of dermal fibrosis. *Faseb J* 2008;22:2214–22.
36 Akhmetshina A, Venalis P, Dees C, *et al*. Treatment with imatinib prevents fibrosis in different preclinical models of systemic sclerosis and induces regression of established fibrosis. *Arthritis Rheum* 2009;60:219–24.
37 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.
38 Pyonteck SM, Akkari L, Schuhmacher AJ, *et al*. CSF-1R inhibition alters macrophage polarization and blocks glioma progression. *Nat Med* 2013;19:1264–72.
39 Goldthorpe H, Jiang JY, Taha M, *et al*. Occlusive lung arterial lesions in endothelial-targeted, fas-induced apoptosis transgenic mice. *Am J Respir Cell Mol Biol* 2015;53:712–8.
40 Dees C, Schlottmann I, Funke R, *et al*. The wnt antagonists DKK1 and SFRP1 are downregulated by promoter Hypermethylation in systemic sclerosis. *Ann Rheum Dis* 2014;73:1232–9.
41 Moldovan L, Moldovan NI. Role of monocytes and macrophages in angiogenesis. *Exs* 2005;94:127–46.
42 Riabov V, Gudima A, Wang N, *et al*. Role of tumor associated macrophages in tumor angiogenesis and lymphangiogenesis. *Front Physiol* 2014;5:75.

43 Okazaki T, Ebihara S, Takahashi H, *et al*. Macrophage colony-stimulating factor induces vascular endothelial growth factor production in skeletal muscle and promotes tumor angiogenesis. *J Immunol* 2005;174:7531–8.

CONCISE REPORT

Interleukin-6 blockade raises LDL via reduced catabolism rather than via increased synthesis: a cytokine-specific mechanism for cholesterol changes in rheumatoid arthritis

Jamie Robertson,[1] Duncan Porter,[2] Naveed Sattar,[3] Chris J Packard,[3] Muriel Caslake,[3] Iain McInnes,[1] David McCarey[4]

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-2017-211708).

[1]Institute of Infection, Immunity & Inflammation, University of Glasgow, Glasgow, UK
[2]Department of Rheumatology, Gartnavel General Hospital, Glasgow, UK
[3]Institute of Cardiovascular & Medical Sciences, University of Glasgow, Glasgow, UK
[4]Centre for Rheumatic Diseases, Glasgow Royal Infirmary, Glasgow, UK

Correspondence to
Dr Jamie Robertson, Institute of Infection, Immunity & Inflammation, University of Glasgow, Glasgow G12 8TA, UK;
jamie.robertson@glasgow.ac.uk

IMI and DMC contributed equally.

Received 30 April 2017
Revised 22 August 2017
Accepted 8 September 2017
Published Online First
15 September 2017

ABSTRACT

Objectives Patients with rheumatoid arthritis (RA) have reduced serum low-density lipoprotein cholesterol (LDL-c), which increases following therapeutic IL-6 blockade. We aimed to define the metabolic pathways underlying these lipid changes.

Methods In the KALIBRA study, lipoprotein kinetic studies were performed on 11 patients with severe active RA at baseline and following three intravenous infusions of the IL-6R blocker tocilizumab. The primary outcome measure was the fractional catabolic rate (FCR) of LDL.

Results Serum total cholesterol (4.8 vs 5.7 mmol/L, p=0.003), LDL-c (2.9 vs 3.4 mmol/L, p=0.014) and high-density lipoprotein cholesterol (1.23 vs 1.52 mmol/L, p=0.006) increased following tocilizumab therapy. The LDL FCR fell from a state of hypercatabolism to a value approximating that of the normal population (0.53 vs 0.27 pools/day, p=0.006). Changes in FCR correlated tightly with changes in serum LDL-c and C-reactive protein but not Clinical Disease Activity Index.

Conclusions Patients with RA have low serum LDL-c due to hypercatabolism of LDL particles. IL-6 blockade normalises this catabolism in a manner associating with the acute phase response (and thus hepatic IL-6 signalling) but not with RA disease activity as measured clinically. We demonstrate that IL-6 is one of the key drivers of inflammation-driven dyslipidaemia.

INTRODUCTION

Patients with rheumatoid arthritis (RA) have serum low-density lipoprotein cholesterol (LDL-c) levels lower than those of age-matched and sex-matched controls,[1–3] despite also having an approximately 50% greater risk of developing cardiovascular disease.[4 5] Conversely, increases in LDL-c or LDL particle numbers have also been observed following treatment of RA with the interleukin-6 (IL-6) receptor blocker tocilizumab[6–8] and the Janus kinase (JAK) inhibitors.[9] The mechanisms underlying this so-called 'lipid-paradox', and the influence of therapy-driven LDL-c changes on cardiovascular risk, remain only partially understood.[10]

Reduced serum LDL-c might be due to reduced LDL synthesis (predominantly from lipolysis of very low-density lipoprotein (VLDL) and intermediate-density lipoprotein (IDL)) or increased LDL turnover. Hypercatabolism of LDL has previously been identified in patients with hypertriglyceridaemia, with catabolic rates normalising following reduction of serum triglyceride (TG) with fibrates.[11]

Tocilizumab is an established and effective treatment for RA which has been shown to increase LDL-c by up to 20% on average.[7] We therefore hypothesised that tocilizumab would reduce LDL catabolism and thus increase serum LDL-c in patients with active RA.

METHODS

In the KALIBRA study (**K**inetics of the **A**poB-containing **L**ipoproteins in **I**L-6 **B**lockade for **RA**) we performed kinetic studies on patients with active RA before and after 10 weeks' treatment with tocilizumab. The primary outcome measure was the change in fractional catabolic rate (FCR) of LDL-associated apolipoprotein (Apo) B).

Recruitment

Subjects were recruited from rheumatology clinics in Glasgow, UK. Patients were provided with information on the study for at least 48 hours before a screening visit where written, informed consent was provided. All further assessments took place at the Clinical Research Facility of the Western Infirmary Glasgow. Subjects met the following inclusion criteria: RA (2010 ACR criteria); DAS28 ≥5.1; failure of two conventional disease-modifying antirheumatic drugs, including methotrexate; and suitability for tocilizumab therapy. Exclusion criteria included: familial dyslipidaemia; ApoE 2/2 homozygosity; diabetes mellitus; use of lipid-lowering therapeutics; fasting total cholesterol (TC) ≥6.5 mmol/L; fasting TG ≥3 mmol/L; pregnancy or untreated hypothyroidism. Use of oral steroid was permitted at a steady dose, but parenteral corticosteroid was prohibited.

Ethical approval for the KALIBRA study was granted locally by the West of Scotland Research Ethics Committee.

Kinetic studies

Each subject underwent a kinetic study at baseline, followed by at least 10 weeks (ie, three infusions) of tocilizumab 8 mg/kg intravenously, and a subsequent 'on-treatment' kinetic study.

To cite: Robertson J, Porter D, Sattar N, *et al*. *Ann Rheum Dis* 2017;**76**:1949–1952.

Table 1 Demographic data of KALIBRA subjects at baseline (n=12)

	Mean (%)
Age	49.9
Sex (F)	10 (83)
RF/ACPA +	10 (83)
Methotrexate	4 (33)
Other DMARD	8 (67)
Prednisolone	2 (17)
NSAID	8 (67)
Previous biological therapy	3 (25)

ACPA, anti-citrullinated peptide antibody; DMARD, disease-modifying antirheumatic drug; NSAID, non-steroidal anti-inflammatory drug; RF, rheumatoid factor.

In each kinetic study, a tracer in the form of the stable isotope d3-leucine at 10 mg/kg body weight bolus was administered. Leucine is taken up by hepatocytes and incorporated into ApoB. Fasting blood samples were obtained at 23 timepoints over 96 hours. Density-gradient ultracentrifugation, precipitation and gas chromatography mass spectrometry could then be used to determine the tracer/tracee ratio in all the ApoB-containing lipoproteins at each timepoint. A 15-compartment mathematical model was used to analyse the the lipolytic pathway through VLDL1, VLDL2, IDL and LDL. SAAM software was then employed to calculate the production rate (PR) and FCR of LDL.

Sample handling and processing

At each kinetic study, blood samples were obtained for beta-quantification of lipids; CRP and erythrocyte sedimentation rate (ESR); lipoprotein (a) (Lp(a)); insulin; proprotein convertase subtilisin kexin type 9 (PCSK9) levels; apolipoproteins by immunoturbidimetry; and activity of cholesteryl ester transfer protein (CETP) and heparin-inducible lipases. DAS28 and Clinical Disease Activity Index (CDAI) were assessed by a clinical research fellow in rheumatology.

Statistical analyses

Statistical analysis was performed at the Robertson Centre for Biostatistics, University of Glasgow. Pretreatment and post-treatment analyses were performed using paired t-test or Wilcoxon matched-pairs test. Correlations were performed using Spearman's *r*, due to the high prevalence of non-parametric data.

In normal subjects, the FCR of LDL-associated ApoB is around 0.3 pools per day. In those with increased catabolism due to hypertriglyceridaemia in a previous study, the FCR is around 0.5 pools per day; this corresponded to a large decrease in LDL-c of around 1.5 mmol/L.[12] In contrast, ciprofibrate has been shown to increase LDL FCR from 0.32 to 0.38 pools per day, with a 22% fall in LDL-c.[13] This is similar to LDL-c changes seen following tocilizumab therapy, and reflects a biologically significant change in FCR. Using these figures, we used a conservative change in FCR of around 0.05 pools per day to determine power. A sample size of 15 subjects allows us to detect a difference in FCR of 0.05 with SD 0.05 at 90% power and alpha error at 5%. This sample size is typical for kinetic trials of this type.

Reagent supply

D3-Leucine was prepared by Tayside Pharmaceuticals Ltd, Dundee, UK. Tocilizumab for this study was graciously provided by Roche Products Ltd.

RESULTS

Demographics

Twelve subjects were recruited (table 1). One subject withdrew before their second kinetic study, leaving useable data for 11 subjects.

Clinical response

At baseline, mean DAS28-CRP and CDAI were 5.16 and 29.9, respectively. Changes in parameters of disease activity are summarised in the online supplementary figure S1. After treatment, seven subjects were in DAS28 remission, while one met the criteria for CDAI remission.

Serum cholesterol

Beta-quantification of serum lipids is shown in the online supplementary table S1. Elevations were observed in TC, LDL-c and high-density lipoprotein cholesterol (HDL-c). No change was seen in the TC/HDL-c ratio.

LDL kinetics

Production and catabolic rates for LDL are displayed in figure 1. Median (IQR) FCR, the primary outcome measure for the study, fell from 0.53 (0.32–0.68) pools/day to 0.27 (0.19–0.37) pools/day (p=0.002), with median change from baseline of −30%. LDL PR also fell significantly, with reduced ApoB transfer from VLDL-2 through IDL to LDL.

LDL FCR associated strongly with LDL-c both at baseline and in degree of change from baseline (figure 2A). At baseline, FCR correlated with CRP (r=0.74, p=0.012) and showed a non-significant trend to association with ESR (r=0.54, p=0.091). However, FCR did not associate at all with clinical assessment of disease activity, as measured by CDAI (r=0.04, p=0.91) (figure 2B). Similar, non-significant trends were seen in degree

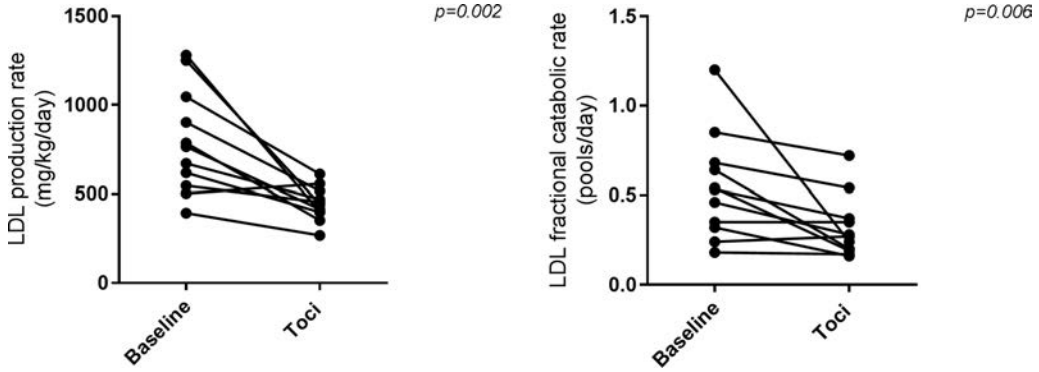


Figure 1 Changes in LDL production rate and FCR after treatment with tocilizumab (N=11). P value generated by Wilcoxon matched-pairs test. FCR, fractional catabolic rate; LDL, low-density lipoprotein.

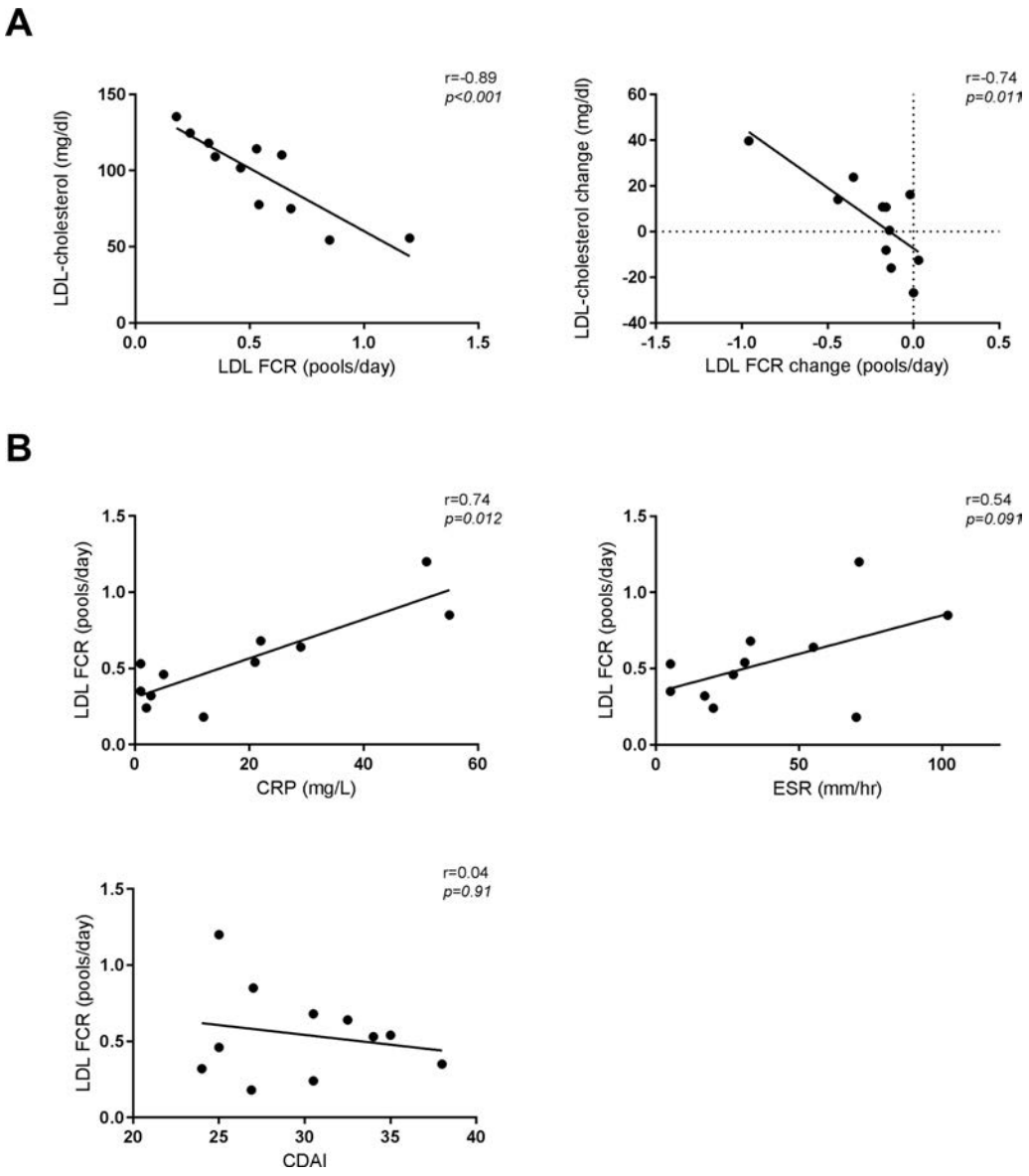


Figure 2 (A) Correlations of LDL FCR with LDL-c at baseline (left) and in degree of change (right). (B) Correlations of LDL FCR with markers of RA disease activity. N=11. R and p values calculated by Spearman's *r* test. CDAI, Clinical Disease Activity Index; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; FCR, fractional catabolic rate; LDL-c, low-density lipoprotein cholesterol; RA, rheumatoid arthritis.

of change for FCR versus CRP (r=0.46, p=0.15), ESR (r=0.30, p=0.37) and CDAI (r=−0.37, p=0.26)

Secondary outcomes

Serum Lp(a) fell, while ApoAI, ApoAII and ApoB increased. Activity levels of CETP, lipoprotein lipase and hepatic lipase, and serum levels of insulin or PCSK9, did not change (online supplementary table S2).

Safety

Two serious adverse events were reported. One patient developed a subcutaneous abscess; this was successfully treated, and the patient resumed therapy and completed the study. Another patient developed paronychia and withdrew from the study.

DISCUSSION

KALIBRA is the first study to demonstrate increased catabolism of LDL particles in active RA. This hypercatabolism correlates tightly with serum LDL-c and acute-phase reactants but not with clinical measures of RA activity, and normalises following IL-6 blockade with tocilizumab. Tocilizumab also reduced LDL production, though this appears to be of negligible biochemical significance as LDL-c is ultimately increased.

These findings allow us to draw some conclusions. First, IL-6 appears to be a key driver of LDL change in RA. We surmise this from tocilizumab's mechanism of action, but also from the association of LDL kinetics with the acute-phase response (driven by IL-6) rather than clinical disease activity; hence, hepatic IL-6 signalling, rather than synovitic bulk, might lie behind LDL changes. Finally, these results are reassuring in terms of the safety of IL-6 blockade, as normalisation of pathological LDL hypercatabolism seems unlikely to drive new atheroma formation.

Charles-Schoeman *et al* previously performed lipid kinetic studies in 33 RA subjects before and after 6 weeks of treatment with tofacitinib, a Janus kinase (JAK) inhibitor, and 31 healthy volunteers.[14] This cohort had increased catabolism of cholesterol ester (but not LDL-associated ApoB) at baseline, which was reduced by tofacitinib. The authors hypothesised decreasing

process of reduced hepatic HDL clearance, with LDL then gaining cholesterol via CETP, reflecting a potentially different mechanism to our own observations.

We believe that KALIBRA holds some advantages over this earlier work. First, using tocilizumab allows us to identify IL-6 as the key molecule behind our results. Tofacitinib, as an inhibitor of multiple JAKs (and thus many upstream cytokines), cannot provide such precise mechanistic information. Second, all our patients had severe active RA with DAS28 ≥5.1. This makes our results more relevant to clinical practice, and gives greater scope for detectable inflammation-driven pathology (and subsequent detectable reversal of that pathology). Charles-Schoeman did not provide values for DAS28, CDAI, ESR or CRP, though the inclusion criteria seem to indicate lower disease activity than our own cohort. Third, our collection of clinical data pretreatment and post-treatment allowed us to ascertain clinical response to the drug and correlate this with metabolic changes.

Our study has some important limitations. The small sample size, while in keeping with previous such kinetic studies, limited our ability to analyse secondary outcome data. Local technical limitations precluded a control group; however, the magnitude of kinetic changes observed are close to previously-described population values of hypercatabolism (at baseline) and normal metabolism (post-treatment).

In conclusion, we demonstrate for the first time that serum LDL-c levels are reduced in active RA due to abnormally elevated LDL catabolism. Additionally, we show that this hypercatabolic state is exquisitely linked to IL-6 signalling, and normalises following therapeutic IL-6 blockade regardless of any clinical reduction in disease activity.

Contributors NS and IM conceived the study. IM, NS and CP wrote the study protocol. DM acted as principle investigator for the study. JR drafted and revised the manuscript. All authors reviewed the draft manuscript and provided feedback for the final version.

Funding This study was sponsored by Roche Products Ltd under contract with NHS Greater Glasgow & Clyde and the University of Glasgow. Funding for a clinical research fellow was provided through the University of Glasgow via an NHS endowment fund. Roche Products Ltd had no role in study design, patient recruitment, sample processing, data interpretation or the preparation of this manuscript.

Competing interests IM, NS, DM, CP and DP have received honoraria from or provided consultancy services for Roche / Chugai. JR has received personal fees from Janssen outside the submitted work. MC reports no relevant conflicts of interest.

Patient consent Consent for publication of results was included in the consent form signed by all subjects, as approved by West of Scotland Research Ethics Committee.

Ethics approval West of Scotland Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Full kinetic data for apoB-containing lipoproteins (VLDL-1, VLDL-2, IDL and LDL) may be available on discussion with study authors.

REFERENCES

1 Yoo WH. Dyslipoproteinemia in patients with active rheumatoid arthritis: effects of disease activity, sex, and menopausal status on lipid profiles. *J Rheumatol* 2004;31:1746–53.
2 Chung CP, Oeser A, Raggi P, *et al*. Lipoprotein subclasses determined by nuclear magnetic resonance spectroscopy and coronary atherosclerosis in patients with rheumatoid arthritis. *J Rheumatol* 2010;37:1633–8.
3 Peters MJ, Voskuyl AE, Sattar N, *et al*. The interplay between inflammation, lipids and cardiovascular risk in rheumatoid arthritis: why ratios may be better. *Int J Clin Pract* 2010;64:1440–3.
4 Aviña-Zubieta JA, Choi HK, Sadatsafavi M, *et al*. Risk of cardiovascular mortality in patients with rheumatoid arthritis: a meta-analysis of observational studies. *Arthritis Rheum* 2008;59:1690–7.
5 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.
6 McInnes IB, Thompson L, Giles JT, *et al*. Effect of interleukin-6 receptor blockade on surrogates of vascular risk in rheumatoid arthritis: MEASURE, a randomised, placebo-controlled study. *Ann Rheum Dis* 2015;74:694–702.
7 Smolen JS, Beaulieu A, Rubbert-Roth A, *et al*. Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. *Lancet* 2008;371:987–97.
8 Gabay C, McInnes IB, Kavanaugh A, *et al*. Comparison of lipid and lipid-associated cardiovascular risk marker changes after treatment with tocilizumab or adalimumab in patients with rheumatoid arthritis. *Ann Rheum Dis* 2016;75:1806–12.
9 Fleischmann R, Kremer J, Cush J, *et al*. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. *N Engl J Med* 2012;367:495–507.
10 Robertson J, Peters MJ, McInnes IB, *et al*. Changes in lipid levels with inflammation and therapy in RA: a maturing paradigm. *Nat Rev Rheumatol* 2013;9:513–23.
11 Slater HR, Packard CJ, Shepherd J. Receptor-independent catabolism of low density lipoprotein. Involvement of the reticuloendothelial system. *J Biol Chem* 1982;257:307–10.
12 Packard CJ, Demant T, Stewart JP, *et al*. Apolipoprotein B metabolism and the distribution of VLDL and LDL subfractions. *J Lipid Res* 2000;41:305–18.
13 Gaw A, Packard CJ, Caslake MJ, *et al*. Effects of ciprofibrate on LDL metabolism in man. *Atherosclerosis* 1994;108:137–48.
14 Charles-Schoeman C, Fleischmann R, Davignon J, *et al*. Potential mechanisms leading to the abnormal lipid profile in patients with rheumatoid arthritis versus healthy volunteers and reversal by tofacitinib. *Arthritis Rheumatol* 2015;67:616–25.

Population-based screening for ACPAs: a step in the pathway to the prevention of rheumatoid arthritis?

Anticitrullinated protein/peptide antibodies (ACPAs), in addition to rheumatoid factor (RF), represent a serological hallmark in the diagnosis of rheumatoid arthritis (RA). In this context and with interest I read the recent article by van Zanten *et al*[1] summarising the results from the Dutch 'Lifelines Study', a large population-based study of 40 132 individuals. In this setting and using an adjusted cut-off for the ACPA assay, the prevalence of ACPAs was 1.0% and associated older age, female gender, smoking, joint complaints, RA and first-degree relatives with rheumatism.

The early identification of patients in the preclinical phase of RA is of high importance as it became evident during the last decade that early intervention can prevent joint damage in patients with RA. Consequently, several ongoing studies are focused on the prevention of RA based on the treatment of individuals at high risk to develop RA.[2] All these prevention trials leverage the concept of the 'window of opportunity' to prevent or delay the clinical ravages and attending healthcare expenditures associated with RA. Although a recent trial was unable to prevent the onset of RA (PRAIRI, unpublished data), the study clearly demonstrated that the treatment in the preclinical phase of RA postponed the development of the disease. Most studies aim to treat patients as early as possible, an approach which seems intuitive. However, it is unclear if treatment too early in the clinical course also leads to failure of response. Most conventional approaches to the prevention of RA are based on the concept of restoring the lost tolerance of the immune system.

PRECLINICAL RA

The risk for the development of RA depends on many factors, which can be divided into two main categories: the modifiable (eg, smoking, behaviour) and the constant risk factors (eg, genetics).[3 4] The preclinical phase of RA may be initiated by modification of the risk profile (eg, smoking) and is characterised by break of tolerance as part of the autoimmune processes, leading to increasing joint inflammation and eventually tissue damage and significant morbidity. The development of RA has primarily been linked to factors that affect the gastrointestinal, the respiratory and reticuloendothelial systems.[4] An additional challenge to identify the right time for treatment is the patient-specific rate of the evolution from preclinical to clinical RA. In some patients, the preclinical phase can take several years, whereas in other patients, this conversion happens in a shorter period of time. Factors contributing, to defining or accelerating this transition are not fully understood.

BIOMARKERS

Based on all the findings about the possible treatment in the preclinical phase of RA, reliable biomarkers are needed to identify patients who are on the trajectory to develop RA. Once a panel of biomarkers has been found and carefully validated, population screening for evidence-based, effective treatment is a conceivable approach in the efforts to prevent RA. In this context, recently, the combination of ACPAs, RF and anticarbamylated peptide (CarP) autoantibodies[5] has been shown to provide a very high odds ratio for RA.[6] Unfortunately, RF and anti-CarP autoantibody data were not available for the population described by van Zanten *et al*.[1] As the Lifelines Study is a prospective longitudinal cohort study with a targeted 30-year follow-up, it might be possible to gain follow-up information on this study population. It will be interesting to see which participants will eventually develop RA. In addition, it would be valuable to test the stored serum samples for as many biomarkers as possible in order to identify patients who develop RA in a certain period of time. Lastly, combining biomarker data with clinical parameters might result in sufficient power to precisely predict the development of RA.[7]

Several studies have now repeatedly shown that ACPAs and other biomarkers (eg, autoantibodies, inflammatory proteins, cytokines and microRNA) can antedate the development of RA by many years.[7 8] Although these data are intriguing, it would be more valuable to have biomarkers that are able to provide insights into the evolution of RA within the next 6–12 months, corresponding to the so-called window of opportunity.[9 10] When it comes to population screening, health economics becomes an important factor in the equation. Since the global healthcare expenditures are constantly increasing and reaching non-sustainable thresholds in many jurisdictions, health economic studies of direct and indirect costs will be needed to demonstrate that investments in screening for pre-RA and early interventions or therapies are not outweighed by poor clinical outcomes. Based on remarkable healthcare expenditures on the management of RA, there is a significant opportunity to achieve meaningful savings.[2 11 12]

CLINICAL PREVENTION TRIALS

The design of the different RA prevention trials is rather diverse in terms of the treatment used and the inclusion criteria for the individuals at risk to be treated.[2] Therapeutic agents range from hydroxychloroquine to biologics (eg, tumour necrosis factor-alpha inhibitors). Different treatments surely require also different timings in the preclinical phase of the disease. Lastly, information as to who will develop RA in the distant future might be more useful in case preventive drugs (such as vaccines) become available.

Michael Mahler

Correspondence to Dr Michael Mahler, Research, Inova Diagnostics, 9900 Old Grove Road, San Diego, CA 92131, USA; mmahler@inovadx.com

Competing interests Employee of Inova Diagnostics selling autoantibody assays.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Mahler M. *Ann Rheum Dis* 2017;**76**:e42.

Accepted 2 February 2017
Published Online First 8 March 2017

► http://dx.doi. org/10.1136/annrheumdis-2017-211256

Ann Rheum Dis 2017;**76**:e42. doi:10.1136/annrheumdis-2017-211249

REFERENCES

1 van Zanten A, Arends S, Roozendaal C, *et al*. Presence of anticitrullinated protein antibodies in a large population-based cohort from the Netherlands. *Ann Rheum Dis* 2017; 0: 1–7.
2 Finckh A, Escher M, Liang MH, *et al*. Preventive treatments for rheumatoid arthritis: issues regarding patient preferences. *Curr Rheumatol Rep* 2016;18:51.

3 Deane KD. Preclinical rheumatoid arthritis (autoantibodies): an updated review. *Curr Rheumatol Rep* 2014;16:419.
4 Deane KD, El-Gabalawy H. Pathogenesis and prevention of rheumatic disease: focus on preclinical RA and SLE. *Nat Rev Rheumatol* 2014;10:212–28.
5 Trouw LA, Mahler M. Closing the serological gap: promising novel biomarkers for the early diagnosis of rheumatoid arthritis. *Autoimmun Rev* 2012;12:318–22.
6 Shi J, van Steenbergen HW, van Nies JA, *et al*. The specificity of anti-carbamylated protein antibodies for rheumatoid arthritis in a setting of early arthritis. *Arthritis Res Ther* 2015;17:339.
7 Rakieh C, Nam JL, Hunt L, *et al*. Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study. *Ann Rheum Dis* 2015;74:1659–66.
8 Kokkonen H, Mullazehi M, Berglin E, *et al*. Antibodies of IgG, IgA and IgM isotypes against cyclic citrullinated peptide precede the development of rheumatoid arthritis. *Arthritis Res Ther* 2011;13:R13.
9 van Nies JA, Krabben A, Schoones JW, *et al*. What is the evidence for the presence of a therapeutic window of opportunity in rheumatoid arthritis? A systematic literature review. *Ann Rheum Dis* 2014;73:861–70.
10 Mankia K, Emery P. A new window of opportunity in rheumatoid arthritis: targeting at-risk individuals. *Curr Opin Rheumatol* 2016;28:260–6.
11 Anaya JM, Duarte-Rey C, Sarmiento-Monroy JC, *et al*. Personalized medicine. Closing the gap between knowledge and clinical practice. *Autoimmun Rev* 2016; 15:833–42.
12 Finckh A, Deane KD. Prevention of rheumatic diseases: strategies, caveats, and future directions. *Rheum Dis Clin North Am* 2014;40:771–85.

Identification of Lifelines participants at high risk for development of rheumatoid arthritis

We would like to thank Michael Mahler[1] for his letter, 'Population-based screening for anticitrullinated protein antibodies (ACPA): A step in the pathway to the prevention of rheumatoid arthritis?' In his letter, Dr Mahler makes a plea to analyse and define reliable biomarkers to identify subjects who are on the trajectory to develop rheumatoid arthritis (RA). We agree with this notion of Dr Mahler as current efforts are increasingly focusing on the possibility to install early preclinical treatment to prevent progression to RA. For example, several randomised placebo-controlled trials (RCTs) have been initiated recently. These clinical trials will investigate the therapeutic potential of several immunomodulatory agents such as rituximab (PRAIRI study: NTR No. 1969), abatacept (APIPPRA study: ISRCTN No. 46017566 and ARIAA study: EudraCT No. 2014-000555-93), hydroxychloroquine (StopRA trial: NCT No. 02603146) and methotrexate (TREAT EARLIER: NTR No. 4853) in individuals at risk of RA. A study evaluating the immunomodulatory effect of atorvastatin in patients with seropositive arthralgia has also been initiated (STAPRA study: NTR No. 22389).

Our publication contributes to these and other efforts as it describes the presence of a prominent biomarker for RA, ACPAs, in a large population-based study and thereby aids to the development of predictive algorithms identifying individuals at risk for RA development. As outlined in this publication, we determined the prevalence of ACPA positivity and its association with known RA risk factors in 40 136 participants from the Lifelines cohort.[2] Lifelines is a large ongoing prospective population-based cohort study and biobank to investigate the interaction between environmental and genetic factors in the development of chronic diseases. The standardised protocol includes physical examination, extensive questionnaires, and biobanking of serum, plasma, urine and DNA. Participants will be followed up according to a fixed protocol for at least 30 years.[3 4] Within the existing infrastructure, additional collection of data and additional studies in already biobanked and future samples can be performed. For example, the composition of the microbiome is currently being analysed in over 900 subjects participating in the Lifelines deep cohort,[5] eventually facilitating the contribution of the microbiome composition to the risk to develop RA. Likewise, genetics such as whole genome sequencing including human leukocyte antigen (sub-) typing can be performed, thereby allowing the analyses of the contribution of the microbiome in the susceptible genetic background to the development of autoimmunity.[6] Other described predictive markers like the acute-phase reactants erythrocyte sedimentation rate and C reactive protein and additional markers can be incorporated in the development of prediction models. Indeed, recent insight into the specificity of the combined presence of ACPA, rheumatoid factor and anticarbamylated protein antibodies provides interesting possibilities in further narrowing down persons at risk.[7 8]

Overall, Lifelines will give us the unique opportunity to gain follow-up information on our study population. With funding, it will be possible to measure additional biomarkers as indicated above. Therefore, this cohort is well suited for subsequent analyses on other biomarkers and risk factors, as also advocated by Dr Mahler.

Our future goal is to build a prediction model that can distinguish participants at high risk from those at low risk to develop RA. Such a model is relevant because both animal data and clinical data suggest that prevention of RA may be possible and early treatment aiming at drug-free remission is possible.[9 10]

The performance of RCTs on preventive treatment strategies in individuals at risk for RA is challenging. The development of reliable assays and predictive algorithms to define a population with high enough at risk is crucial to allow the design of effective preventive therapy aiming to restore immune homeostasis with a great specificity, low toxicity and long-term effectiveness.

S Arends,[1] L A Trouw, R E M Toes,[2] A van Zanten,[2,1] C Roozendaal,[3] P C Limburg, H Bootsma,[3,1] E Brouwer[1]

[1]Department of Rheumatology and Clinical Immunology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[2]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[3]Department of Laboratory Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Correspondence to Dr E Brouwer, Department of Rheumatology and Clinical Immunology, University Medical Center Groningen, University of Groningen, Groningen 9713GZ, The Netherlands; e.brouwer@umcg.nl

Competing interests None declared.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Arends S, Trouw LA, Toes REM, *et al*. *Ann Rheum Dis* 2017;**76**:e43.

Accepted 15 February 2017
Published Online First 8 March 2017

▶ http://dx.doi.org/10.1136/annrheumdis-2017-211249

Ann Rheum Dis 2017;**76**:e43. doi:10.1136/annrheumdis-2017-211256

REFERENCES

1 Mahler M. Population based screening for ACPA: a step in the pathway to the prevention of rheumatoid arthritis? *Ann Rheum Dis* 2017;76:e42.
2 Van Zanten A, Arends S, Roozendaal C, *et al*. Presence of anticitrullinated protein antibodies in a large population-based cohort from the Netherlands. *Ann Rheum Dis* 2017. doi:10.1136/annrheumdis-2016-209991 [Epub ahead of print: 2 Jan 2017].
3 Scholtens S, Smidt N, Swertz MA, *et al*. Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol* 2015;44:1172–80.
4 Klijs B, Scholtens S, Mandemakers JJ, *et al*. Representativeness of the LifeLines Cohort Study. *PLoS ONE* 2015;10:e0137203.
5 Bonder MJ, Kurilshikov A, Tigchelaar EF, *et al*. The effect of host genetics on the gut microbiome. *Nat Genet* 2016;48:1407–12.
6 van Heemst J, Jansen DTSL, Polydorides S, *et al*. Crossreactivity to vinculin and microbes provides a molecular basis for HLA-based protection against rheumatoid arthritis. *Nat Commun* 2015;6:6681.
7 Shi J, van Steenbergen HW, van Nies JA, *et al*. The specificity of anti-carbamylated protein antibodies for rheumatoid arthritis in a setting of early arthritis. *Arthritis Res Ther* 2015;17:339.
8 Koppejan H, Trouw LA, Sokolove J, *et al*. Role of anti-carbamylated protein antibodies compared to anti-citrullinated protein antibodies in indigenous North Americans with rheumatoid arthritis, their first-degree relatives, and healthy controls. *Arthritis Res Ther* 2016;68:2090–8.
9 Dekkers JS, Schoones JW, Huizinga TW, *et al*. Possibilities for preventive treatment in rheumatoid arthritis? Lessons from experimental animal models of arthritis: a systematic literature review and meta-analysis. *Ann Rheum Dis* 2017;76:458–67.
10 Ajeganova S, van Steenbergen HW, van Nies JA, *et al*. Disease-modifying antirheumatic drug-free sustained remission in rheumatoid arthritis: an increasingly achievable outcome with subsidence of disease symptoms. *Ann Rheum Dis* 2016;75:867–73.

Prospective MRI score to predict negative EULAR response in patients with rheumatoid arthritis (RA) before therapy-escalation to a biological therapy

Dear Editor

We read with great interest the article by Baker *et al*[1] who showed that early MRI measures independently predict erosive progression on X-ray and MRI after 1 and 2 years in therapy-naive patients with rheumatoid arthritis (RA) from the randomised-controlled GO-BEFORE trial. Due to these findings, we re-evaluated MRI data from the German REMISSION-PLUS Cohort[2 3] at our centre to verify if a MRI score may predict negative response in patients with RA before therapy-escalation to a biological therapy. MRI was performed in 257 patients before therapy-escalation (T0) and after 12 months (T1) and analysed by using the Outcome Measures in Rheumatology (OMERACT) rheumatoid arthritis MRI score (RAMRIS). In addition, clinical and laboratory parameters (Disease Activity Score 28 (DAS-28) and C-reactive protein (CRP)) were collected for each visit. Logistic regression combining clinical and MRI parameters was performed resulting in a combination of the patients' age and the RAMRIS-T0 performing best for prediction of non-response. Bootstrapping with 5000 resamples was performed to estimate the accuracy of the model.

Of the patients included, 29 were escalated to a biological therapy (20 women, median age 57 years (IQR 46–65), 95% anti-tumour necrosis factor (TNF)-alpha therapy). Poor responders (n=5) and responders (n=24) had a mean RAMRIS-T0 score of 14.4 and 52.0, respectively (Wilcoxon test $p<0.01$). High RAMRIS score showed a trend towards a protective effect against non-response (OR 0.90 per RAMRIS point, 95% CI 0.79 to 1.03, $p=0.12$). The strength of the association was stable after adjusting for age, CRP, anti citrullinated peptide antibodies (ACPA)/rheumatoid factor and DAS-28 at baseline. The median area under the curve in the bootstrap analysis was 88.9% with 95% CI 84.0% to 92.8%.

Thus, while Baker *et al* clearly demonstrated that a high inflammatory activity on MRI (ie, RAMRIS) is associated with an unfavourable prognosis (ie, radiographic progression), our observations suggest that this may be overcome by administration of a highly effective therapy, for example, a biologic agent. Indeed, patients with a prognostic unfavourable high RAMRIS were even more likely to respond, making them ideal candidates for these costly drugs.

In summary, both studies emphasise the value of an MRI before therapy initiation or escalation. Hence, further studies are needed to improve our data in established patients with RA before escalating the therapy to biological disease-modifying anti-rheumatic drug (bDMARD).

Philipp Sewerin, Stefan Vordenbaeumen, Ralph Brinks, Benedikt Ostendorf

Department of Policlinic of Rheumatology and Hiller Research Unit Rheumatology University Hospital (UKD), Duesseldorf, Germany

Correspondence to Dr Philipp Sewerin, Department and Hiller Research Unit for Rheumatology, University Hospital Duesseldorf (UKD), Moorenstreet 5, 40225 Duesseldorf, Germany; philipp.sewerin@med.uni-duesseldorf.de

Competing interests None declared.

Provenance and peer review Not commissioned; internally peer reviewed.

CrossMark

To cite Sewerin P, Vordenbaeumen S, Brinks R, *et al*. *Ann Rheum Dis* 2017;**76**:e44.

Received 3 February 2017
Accepted 4 February 2017
Published Online First 10 March 2017

Linked

► http://dx.doi. org/10.1136/annrheumdis-2017-211265

Ann Rheum Dis 2017;**76**:e44. doi:10.1136/annrheumdis-2017-211257

REFERENCES

1 Baker JF, Ostergaard M, Emery P, *et al*. Early MRI measures independently predict 1-year and 2-year radiographic progression in rheumatoid arthritis: secondary analysis from a large clinical trial. *Ann Rheum Dis* 2014;73:1968–74.
2 Ostendorf B, Scherer A, Kellner H, *et al*. Project REMISSION(PLUS): clinical and radiological remission: new treatment goals in the management of rheumatoid arthritis. *Z Rheumatol* 2008;67:707–10, 712–15.
3 Sewerin P, Klein S, Brinks R, *et al*. REMISSIONPLUS, an Initiative to Integrate Modern Imaging into Rheumatologic Care – Review, Appraisal and Outlook: Evaluation of Low-Field MRI Data. *Akt Rheumatol* 2017;42:1–10 (2017, in press).

Is MRI a predictive biomarker for clinical response to biologics in rheumatoid arthritis?

We thank Sewerin *et al*[1] for the data they have provided on the predictive role of MRI for clinical response in rheumatoid arthritis (RA). This study follows on from our previous study demonstrating the predictive value of MRI for radiographic damage progression.[2] This study by Sewerin builds on the evidence for MRI as an imaging biomarker by demonstrating a prediction of clinical response. The investigators used the German REMISSION-PLUS[3] cohort and studied 29 patients who were being escalated to biologic therapy due to inadequate disease control. Clinical European League Against Rheumatism response to the biologic was more likely in those with higher RA MRI scores prior to the escalation of therapy. While these study results need to be replicated in a larger cohort, this study provides initial evidence that MRI measures can help predict who is most likely to benefit from more aggressive interventions.

The concept that MRI more accurately identifies clinically relevant synovitis than clinical assessment is well established.[4] One hypothesis for why MRI might be a useful predictive biomarker for therapeutic response is that some patients with apparently active RA have elevated disease activity measures due to comorbid conditions, rather than active joint inflammation. For example, a recent study showed that obese patients are less likely to reach clinical remission.[5] Those with elevated disease activity measures without objective evidence of inflammatory disease would be very unlikely to improve with more aggressive treatment of the RA. In contrast, those with greater MRI-detected activity might be expected to have a greater proportion of their clinical disease activity explained by active RA.

Limitations of this study are the small sample size and lack of a more detailed characterisation of the study population. However, this study begins to answer an important question in RA, namely—can MRI help rheumatologists make more accurate decisions about escalation of therapy? Given that escalation to biologic therapy involves increased risk of side effects and cost, biomarkers that define both cases at greatest risk of joint destruction and those most likely to benefit clinically are of major interest. MRI, despite being expensive, is likely to be cost-effective in circumstances when its use prevents unnecessary or inappropriate use of much more expensive and long-term therapies.

Joshua F Baker,[1,2,3] **Mikkel Østergaard,**[4,5] **Philip G Conaghan**[6]

[1]Philadelphia Veterans Affairs Medical Center, Philadelphia, Pennsylvania, USA
[2]University of Pennsylvania, School of Medicine, Philadelphia, Pennsylvania, USA
[3]Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[4]Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark
[5]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen Denmark
[6]NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK

Correspondence to Dr Joshua F Baker, Philadelphia Veterans Affairs Medical Center, Philadelphia, PA 19104, USA; bakerjo@uphs.upenn.edu

Competing interests None declared.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Baker JF, Østergaard M, Conaghan PG. *Ann Rheum Dis* 2017;**76**:e45.

Received 7 February 2017
Revised 21 February 2017
Accepted 22 February 2017
Published Online First 10 March 2017

► http://dx.doi.org/10.1136/annrheumdis-2017-211257

Ann Rheum Dis 2017;**76**:e45. doi:10.1136/annrheumdis-2017-211265

REFERENCES

1 Sewerin P, Vordenbaeumen S, Brinks R, *et al*. Prospective MRI-score to predict negative EULAR-response in patients with rheumatoid arthritis (RA) before therapy-escalation to a biological therapy. *Ann Rheum Dis* 2017;76:e44.
2 Baker JF, Ostergaard M, Emery P, *et al*. Early MRI measures independently predict 1-year and 2-year radiographic progression in rheumatoid arthritis: secondary analysis from a large clinical trial. *Ann Rheum Dis* 2014;73:1968–74.
3 Ostendorf B, Scherer A, Kellner H, *et al*. [Project REMISSION(PLUS): clinical and radiological remission : new treatment goals in the management of rheumatoid arthritis]. *Z Rheumatol* 2008;67:707–10, 12–15.
4 Brown AK, Conaghan PG, Karim Z, *et al*. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum* 2008;58:2958–67.
5 Liu Y, Hazlewood GS, Kaplan GG, *et al*. Impact of obesity on remission and disease activity in rheumatoid arthritis: a systematic review and meta-analysis. *Arthritis Care Res (Hoboken)* 2017;69:157–65.

How to diagnose IgG4-related disease

We read with great interest the editorial by Fox and Fox[1] describing the use of serum immunoglobulin G4 (IgG4) concentrations as a marker for IgG4-related disease (IgG4-RD). IgG4-RD is a fascinating clinical entity including a wide variety of diseases, formerly diagnosed as Mikulicz's disease, autoimmune pancreatitis (AIP), interstitial nephritis, prostatitis and retroperitoneal fibrosis.[2 3] However, universal criteria for IgG4-RD have not yet been established at present, making its diagnosis in some patients ambiguous leading to many IgG4-RD mimickers.

A 3-year investigation by the Japanese IgG4 team, organised by the Ministry of Health, Labour and Welfare (MHLW) of Japan, has reached a consensus, in that IgG4-RD can occur in various organs, with clinical symptoms depending on lesion location. Characteristics common to all forms of IgG4-RD include elevated serum IgG4 concentration and tissue infiltration by IgG4-positive plasma cells, accompanied by tissue fibrosis and sclerosis.[2] In 2011, the Japanese IgG4 team published comprehensive diagnostic (CD) criteria for IgG4-RD,[4] with the major characteristics being serum IgG4 concentration >135 mg/dL, the infiltration of >10 IgG4+ cells per high-powered field (HPF) and an IgG4+/IgG+ cell ratio >40%. The cut-off of 135 mg/dL was based on receiver operating characteristic curves and its validity was confirmed in patients with AIP.[5 6] Since then, serum IgG4 levels have been widely used as a reliable criterion for the diagnosis of IgG4-RD.[4 6]

However, Dr Fox mentioned drawbacks of using serum IgG4 levels in diagnosing IgG4-RD, citing studies reporting that the IgG4 cut-off >135 mg/dL had a low sensitivity and specificity for the diagnosis of IgG4-RD.[7 8] As increased serum concentrations of IgG4 have been observed in several diseases with aberrant immunological condition unrelated to IgG4-RD, such as malignant tumours, autoimmune diseases especially rheumatoid arthritis and allergic diseases,[9 10] increased IgG4 concentration is not a specific marker for IgG4-RD. In contrast, recent large cohort studies from the UK, Taiwan and Japan showed that serum IgG4 concentration >135 mg/dL had overall sensitivities of 82.8%, 86% and 88%, respectively, in diagnosing IgG4-RD.[10–12] As no universal criteria for IgG4-RD have been developed to date, three criteria such as international consensus diagnostic criteria for AIP,[13] consensus statement on the pathology[14] and CD criteria[4] have been often used for diagnosis of IgG4-RD. Therefore, the sensitivity and specificity of specific markers may differ among studies that use different diagnostic criteria.

Since this complex multisystem disease represented a single pathogenetic disorder manifesting in a variety of target organs, the diagnosis of IgG4-RD is largely based on biopsy results showing enhanced infiltration by IgG4-positive plasma cells, storiform fibrosis, obliterative phlebitis and moderate eosinophilia, all of which are frequently observed in the affected tissues of these patients.[2 3] A high number of IgG4-positive plasma cells in tissue is a hallmark of IgG4-RD, even when serum IgG4 concentrations are below the cut-off level. The number of IgG4-positive plasma cells differ among organs, and consensus statement on the pathology emphasises tissue IgG4 cell counts in each organ for diagnosis of IgG4-RD.[14] However, these counts should be supplemented by IgG4+/IgG+ plasma cell ratio of more than 40% to distinguish IgG4-RD.[4]

As stated by Fox and Fox,[1] IgG4-RD tends to be both underdiagnosed and overdiagnosed. Underdiagnosis is due to a lack of recognition of this disease, and overdiagnosis results from the well-intentioned enthusiasm of physicians and/or pathologists who recognise IgG4-RD and diagnose similar conditions as IgG4-RD. Therefore, simple and strict criteria are required in the diagnosis of patients with IgG4-RD. In this point, a definite diagnosis of IgG4-RD by CD criteria requires that patients satisfy all three diagnostic characteristics: clinical evidence, high (>135 mg/dL) serum IgG4 and pathological certification (>10 IgG4+ cells/HPF and IgG4+/IgG cell ratio >40%), although some patients may not satisfy these specific serological and/or histopathological criteria because of the difficulty of obtaining biopsies, and therefore cannot be diagnosed with definite IgG4-RD.[4]

To resolve this problem, several Japanese medical societies, including those for gastroenterology, pancreas, biliary tract, rheumatology, ophthalmology and respiratology, have published organ specific criteria for IgG4-RD.[13 15–18] Each criterion contains organ-specific clinical symptom and characteristic radiological findings of IgG4-RD, even with steroidal trial in some

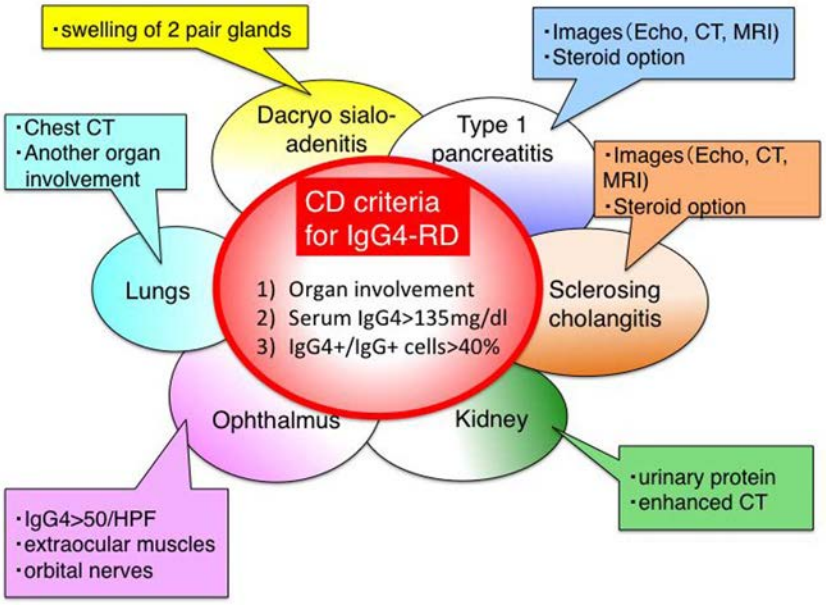


Figure 1 Combination of comprehensive diagnostic (CD) criteria and organ-specific criteria for diagnosing IgG4-related disease (IgG4-RD). HPF, high powered field; IgG4, immunoglobulin G4.

criteria. We recently published a paper describing the optimal method of diagnosing IgG4-RD, based on combinations of CD and organ-specific criteria[19] (figure 1).

None of the diagnostic criteria have been approved to date by the American College of Rheumatology for reasons that include insufficient sensitivity and specificity, implications for billing and reimbursement, healthcare priorities and treatment implications for patients.[20] Physicians in every field of medicine, however, may encounter this new disease in daily practice, and proper diagnostic criteria are required immediately. Therefore, we believe that a careful and intensive judgement using combination of CD and organ-specific criteria is the current best way for diagnosis of IgG4-RD.

Hisanori Umehara,[1] Kazuichi Okazaki,[2] Mitsuhiro Kawano,[3] Tsuneyo Mimori,[4] Tsutomu Chiba[5]

[1]Division of RA and Autoimmune Diseases, Nagahama City Hospital, Nagahama, Shiga, Japan
[2]Division of Gastroenterology and Hepatology, The Third Department of Internal Medicine, Kansai Medical University, Osaka, Japan
[3]Division of Rheumatology, Department of Internal Medicine, Graduate School of Medical Science, Kanazawa University, Ishikawa, Japan
[4]Department of Clinical Immunology, Graduate School of Medicine, Kyoto University, Kyoto, Japan
[5]Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kyoto University, Kyoto, Japan

Correspondence to Professor Hisanori Umehara, Division of Rheumatology and Immunology, Nagahama City Hospital, 313, Oinuicho Nagahama, Shiga 526-0043, Japan; umehara606@gmail.com

Funding This study was supported by Intractable Diseases, the Health and Labour Sciences Research Grants from MHLW, Japan.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Intractable Diseases, the Health and Labour Sciences Research Grants from MHLW, Japan.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Umehara H, Okazaki K, Kawano M, *et al*. *Ann Rheum Dis* 2017;**76**:e46.

Accepted 20 February 2017
Published Online First 10 March 2017

Ann Rheum Dis 2017;**76**:e46. doi:10.1136/annrheumdis-2017-211330

REFERENCES

1 Fox RI, Fox CM. IgG4 levels and plasmablasts as a marker for IgG4-related disease (IgG4-RD). *Ann Rheum Dis* 2015;74:1–3.
2 Umehara H, Okazaki K, Masaki Y, *et al*. A novel clinical entity, IgG4-related disease (IgG4RD): general concept and details. *Mod Rheumatol* 2012;22:1–14.
3 Stone JH, Zen Y, Deshpande V. IgG4-related disease. *N Engl J Med* 2012;366:539–51.
4 Umehara H, Okazaki K, Masaki Y, *et al*. Comprehensive diagnostic criteria for IgG4-related disease (IgG4-RD), 2011. *Mod Rheumatol* 2012;22:21–30.
5 Hamano H, Kawa S, Horiuchi A, *et al*. High serum IgG4 concentrations in patients with sclerosing pancreatitis. *N Engl J Med* 2001;344:732–8.
6 Okazaki K, Kawa S, Kamisawa T, *et al*. Clinical diagnostic criteria of autoimmune pancreatitis: revised proposal. *J Gastroenterol* 2006;41:626–31.
7 Carruthers MN, Khosroshahi A, Augustin T, *et al*. The diagnostic utility of serum IgG4 concentrations in IgG4-related disease. *Ann Rheum Dis* 2015;74:14–18.
8 Wallace ZS, Mattoo H, Carruthers M, *et al*. Plasmablasts as a biomarker for IgG4-related disease, independent of serum IgG4 concentrations. *Ann Rheum Dis* 2015;74:190–5.
9 Su Y, Sun W, Wang C, *et al*. Detection of serum IgG4 levels in patients with IgG4-related disease and other disorders. *PLoS ONE* 2015;10:e0124233.
10 Yu KH, Chan TM, Tsai PH, *et al*. Diagnostic performance of serum IgG4 Levels in patients with IgG4-related disease. *Medicine (Baltimore)* 2015;94:e1707.
11 Inoue D, Yoshida K, Yoneda N, *et al*. IgG4-related disease: dataset of 235 consecutive patients. *Medicine (Baltimore)* 2015;94:e680.
12 Culver EL, Sadler R, Simpson D, *et al*. Elevated serum IgG4 levels in diagnosis, treatment response, organ involvement, and relapse in a prospective IgG4-related disease UK cohort. *Am J Gastroenterol* 2016;111:733–43.
13 Shimosegawa T, Chari ST, Frulloni L, *et al*. International consensus diagnostic criteria for autoimmune pancreatitis: guidelines of the International Association of Pancreatology. *Pancreas* 2011;40:352–8.
14 Deshpande V, Zen Y, Chan JK, *et al*. Consensus statement on the pathology of IgG4-related disease. *Mod Pathol* 2012;25:1181–92.
15 Ohara H, Okazaki K, Tsubouchi H, *et al*. Clinical diagnostic criteria of IgG4-related sclerosing cholangitis 2012. *J Hepatobiliary Pancreat Sci* 2012;19:536–42.
16 Kawano M, Saeki T, Nakashima H, *et al*. Proposal for diagnostic criteria for IgG4-related kidney disease. *Clin Exp Nephrol* 2011;15:615–26.
17 Goto H, Takahira M, Azumi A, Japanese Study Group for IgG4-related Ophtalmic Disease. Diagnostic criteria for IgG4-related ophthalmic disease. *Jpn J Ophthalmol* 2015;59:1–7.
18 Matsui S, Yamamoto H, Minamoto S, *et al*. Proposed diagnostic criteria for IgG4-related respiratory disease. *Respir Investig* 2016;54:130–2.
19 Umehara H, Okazaki K, Nakamura T, *et al*. Current approach to the diagnosis of IgG4-related disease-combination of comprehensive diagnostic and organ-specific criteria. *Mod Rheumatol* 2017. doi: 10.1080/14397595.2017.1290911. [Epub ahead of print 6 Feb 2017].
20 Aggarwal R, Ringold S, Khanna D, *et al*. Distinctions between diagnostic and classification criteria? *Arthritis Care Res (Hoboken)* 2015;67:891–7.

2016 update of the EULAR recommendations for the management of rheumatoid arthritis: a utopia beyond patients in low/middle income countries?

We read with great interest the recently published recommendations by the European League against Rheumatism (EULAR) on the management of rheumatoid arthritis (RA).[1] The EULAR recommendations, although primarily targeted towards European countries, are read and followed across the world including low/middle income nations. Consequently, we were disappointed to note that the updated guidelines recommend the use of biological disease-modifying anti-rheumatic drugs (bDMARDs) or targeted synthetic DMARDs (tsDMARDs) immediately following failure of monotherapy with conventional synthetic DMARDs (csDMARDs) in those patients with poor prognostic factors such as seropositivity for rheumatoid factor (RF) or anticitrullinated peptide antibodies (ACPA), highly active disease or early radiographic joint damage (recommendation number 8).[1] This is in contrast to the 2015 guidelines provided by the American College of Rheumatology (ACR) for the management of RA,[2] which offer the option of either combining csDMARDs or using bDMARDs or tofacitinib (tsDMARD) following failure of methotrexate monotherapy in RA, irrespective of the presence or absence of such poor prognostic indicators. Early use of bDMARDs in the management of RA poses certain specific problems, as discussed below.

Rheumatoid arthritis is one of the most common rheumatic diseases.[3] We exemplify India to provide an estimate of the actual burden of RA in a low/middle income country. The population prevalence of RA in India is 0.75%.[4] According to the 2011 Census of India, with a population of 1.21 billion,[5] an estimated 9 million people could be affected with RA. Approximately 30% patients with RA will respond to methotrexate monotherapy.[6] A vast majority of patients with RA have an adverse prognostic factor in the form of seropositivity for RF or ACPA. Hence, in a country like India, most of 6.3 million patients with RA would require bDMARDs as per current guidelines. The healthcare costs of providing long-term bDMARDs to such a large number of patients, mostly without medical insurance or social security, are beyond the capacity of individual patients or governments of most low/middle income countries. This is an even bigger problem when one considers that there is a paucity of guidelines on when to taper and stop DMARDs, including bDMARDs in RA, as also mentioned in the current EULAR recommendations (recommendations 11 and 12).[1]

With this background, we strongly suggest that the cost-effective strategy of treating RA with a combination of csDMARDs when methotrexate monotherapy fails should not be ignored, despite the presence of poor prognostic factors. The TACIT trial confirmed that use of csDMARDs was non-inferior to the use of anti-tumor necrosis factor (TNF) agents in the management of RA, but associated with substantially lesser costs.[7] It is pertinent to note that most of the trials on bDMARDs in RA, which established their utility for this indication, did so with a combination of bDMARDs and methotrexate.[8] This is emphasised in the current EULAR guidelines which recommend the addition of methotrexate or other csDMARD to bDMARD or tsDMARD in phase II of the treatment strategy (recommendation 9).[1] This brings forth an interesting conundrum, that is, how much of the disease-modifying effect of the bDMARDs was attributable to itself vis-à-vis methotrexate? For example, a closer look at the PREMIER study[9] shows that outcomes at 2 years in terms of the proportion of patients attaining ACR 20, ACR 50 and ACR 70 responses were numerically better or equal for methotrexate monotherapy when compared with adalimumab monotherapy. Two excellent meta-analyses by Graudal *et al*[10 11] reaffirm that the use of bDMARDs is associated with earlier attainment of ACR 50 and ACR 70 responses and numerically lesser radiographic progression of RA in the first 2 years. However, the difference disappears at 2 years of therapy. Moreover, the use of csDMARDs in combination is associated with significantly lesser costs.

Nevertheless, it cannot be denied that the advent of bDMARDs has revolutionised the management of RA in the modern era. However, this must be weighed against the marked immunosuppressive state resulting from the use of bDMARDs, which is a major concern in low/middle income countries wherein infections like tuberculosis are endemic. Use of anti-TNF bDMARDs has been reported to cause infections like leprosy in regions of the world where this disease was not believed to exist like the USA, in a patient who never reported travelling outside this geographical region.[12] This suggests that the use of bDMARDs should be undertaken with due caution under all circumstances.

To conclude, we suggest that combination of csDMARDs should still be considered a viable alternative to bDMARDs or tsDMARDs in patients with RA failing initial monotherapy with methotrexate under most circumstances. A strategy of using a combination of csDMARDs upfront in patients with poor prognostic factors, as suggested by the previous ACR recommendations,[13] may be more reasonable in resource-constrained scenarios. Such patients who fail csDMARDs in combination at 3–6 months should be considered for bDMARDs or tsDMARDs. This might help rationalise the economic burden due to bDMARDs or tsDMARDs, while not depriving the appropriate patient of timely treatment with these drugs. The enthusiasm of using bDMARDs upfront should be tempered with pragmatism and caution given lack of definitive evidence of superiority to csDMARDs in combination, significantly higher costs and risk of infections, especially in low/middle income countries.

Durga Prasanna Misra,[1] Vikas Agarwal,[1] Aman Sharma,[2] Anupam Wakhlu,[3] Vir Singh Negi[4]

[1]Department of Clinical Immunology, Sanjay Gandhi Postgraduate Institute of Medical Sciences (SGPGIMS), Lucknow, India
[2]Cinical Immunology and Rheumatology Services, Department of Internal Medicine, Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh, India
[3]Department of Rheumatology, King George's Medical University, Lucknow, India
[4]Department of Clinical Immunology, Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India

Correspondence to Dr Durga Prasanna Misra, Department of Clinical Immunology, C block, 2nd floor, Sanjay Gandhi Postgraduate Institute of Medical Sciences (SGPGIMS), Rae Bareily Road, Lucknow-226014, India; durgapmisra@gmail.com, dpmisra@sgpgi.ac.in

Contributors DPM, VA, VSN—conception and design of the paper. DPM, VA, AS, AW—acquisition of data, analysis and interpretation of data. DPM, VA—drafting the article. AS, AW, VSN—revising it critically for important intellectual content. DPM, VA, AS, AW, VSN—final approval of the version to be submitted.

Competing interests None declared.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Misra DP, Agarwal V, Sharma A, *et al*. *Ann Rheum Dis* 2017;**76**:e47.

Received 9 March 2017
Accepted 11 March 2017
Published Online First 27 March 2017

► http://dx.doi. org/10.1136/annrheumdis-2017-211455

Ann Rheum Dis 2017;**76**:e47. doi:10.1136/annrheumdis-2017-211446

REFERENCES

1 Smolen JS, Landewé R, Bijlsma J, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017. doi: 10.1136/annrheumdis-2016-210715. [Epub ahead of print 6 Mar 2017]
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 Cross M, Smith E, Hoy D, *et al*. The global burden of rheumatoid arthritis: estima from the Global Burden of Disease 2010 study. *Ann Rheum Dis* 2014;73:1316–22.
4 Misra DP, Agarwal V, Negi VS. Rheumatology in India: a Bird's eye view on organization, epidemiology, training programs and publications. *J Korean Med Sci*tes 2016;31:1013–19.
5 http://www.censusindia.gov.in/Census_Data_2001/National_Summary/National_Summary_DataPage.aspx (accessed 8 Mar 2017).
6 Moreland LW, O'Dell JR, Paulus HE, *et al*. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of Early Aggressive Rheumatoid Arthritis Trial. *Arthritis Rheum* 2012;64:2824–35.
7 Scott DL, Ibrahim F, Farewell V, *et al*. Tumour necrosis factor inhibitors versus combination intensive therapy with conventional disease modifying anti-rheumatic drugs in established rheumatoid arthritis: TACIT non-inferiority randomised controlled trial. *BMJ* 2015;350:h1046.
8 Parida JR, Misra DP, Wakhlu A, *et al*. Is non-biological treatment of rheumatoid arthritis as good as biologics? *World J Orthop* 2015;6:278–83.
9 Breedveld FC, Weisman MH, Kavanaugh AF, *et al*. The PREMIER study: A multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. *Arthritis Rheum* 2006;54:26–37.
10 Graudal N, Jürgens G. Similar effects of disease-modifying antirheumatic drugs, glucocorticoids, and biologic agents on radiographic progression in rheumatoid arthritis: meta-analysis of 70 randomized placebo-controlled or drug-controlled studies, including 112 comparisons. *Arthritis Rheum* 2010;62:2852–63.
11 Graudal N, Hubeck-Graudal T, Faurschou M, *et al*. Combination therapy with and without tumor necrosis factor inhibitors in rheumatoid arthritis: a meta-analysis of randomized trials. *Arthritis Care Res (Hoboken)* 2015;67:1487–95.
12 Scollard DM, Joyce MP, Gillis TP. Development of leprosy and type 1 leprosy reactions after treatment with infliximab: a report of 2 cases. *Clin Infect Dis* 2006;43:e19–22.
13 Singh JA, Furst DE, Bharat A, *et al*. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2012;64:625–39.

Response to: '2016 update of the EULAR recommendations for the management of rheumatoid arthritis: no utopia for patients in low/middle-income countries?' by Misra *et al*

We appreciate the comments provided by Misra *et al*[1] on the 2016 update of the European League Against Rheumatism (EULAR) recommendations for the management of rheumatoid arthritis (RA).[2] They raise a very good point by stating that the use of biologic (b) and targeted synthetic (ts) disease-modifying antirheumatic drugs (DMARDs) in countries like India and many others, is limited by unjustifiable drug prices that are the consequence of various policies by drug manufacturers. Therefore, highly effective drugs are not widely accessible to patients in many countries. While EULAR is highly concerned by this reality, the recommendations are not primarily meant to improve accessibility of drugs worldwide. The EULAR recommendations provide what is regarded to be the optimal therapeutic approach according to the evidence derived from systematic literature reviews (SLRs)[3–5] and expert opinion. If the best therapeutic approach may not be feasible in some countries, alternatives have to be sought. Of course such alternatives would be to switch to another conventional synthetic DMARD (csDMARD) or to add a csDMARD, rather than to add a bDMARD or tsDMARD. Misra *et al* try to imply, though, that such an alternative is as efficacious as adding a bDMARD, and this is a comment that the task force did not agree with.

Misra *et al* refer to the tumour necrosis factor inhibitors against combination intensive therapy (TACIT) trial.[6] However, they did not notice that over the first 6 months the disease activity score 28 (DAS28) continued to be quite high in the triple therapy group compared with the antitumour necrosis factor (anti-TNF) group, suggesting that the treatment target, remission or at least low disease activity, which is not reported, was missed in a vast majority of the patients. Only after a TNF-blocker had been added, clinical disease activity decreased to the levels seen with bDMARD therapy already long before. These data confirm the notion provided by Kiely *et al*, another British group, that the use of additional csDMARDs when methotrexate (MTX) has failed does not convey much benefit.[7] On the other hand, the CareRA trial clearly showed that about three of four patients treated with MTX monotherapy plus glucocorticoids (GC) achieved a DAS28-C-reactive protein<2.6; this is the treatment recommended in the EULAR document. It is therefore unclear, why Misra *et al* call the achievement of 75% excellent outcomes already on application of phase I of the EULAR recommendations a utopia.

Misra *et al* then refer to papers by Graudal *et al* but they fail to refer to the methodological criticism on at least one of these papers[8] and also do not appreciate a Cochrane review revealing that combination therapy is not more beneficial than monotherapy.[9] Finally, by pointing to the data from the PREMIER trial, they just reiterate what the task force has stated regarding combinations of csDMARDs with bDMARDs—many more data also support this decision of the EULAR task force.

Misra *et al* suggest that after failure of MTX triple csDMARD therapy should be commenced. While this may be a good option in certain situations, Misra *et al* apparently do not appreciate that in the Behandel Strategieen trial switching to another monotherapy was as efficacious as step-up combination therapy.[10] The higher rate of adverse events seen with csDMARD combinations versus csDMARD monotherapy does not speak for using csDMARD combinations.[6 11]

We are convinced that Misra *et al* will agree on today's therapeutic target of at least low disease activity in patients with established disease.[12] To this end, they should be reminded that in the Rheumatoid Arthritis: Comparison of Active Therapies trial of patients who failed MTX therapy ACR70 response rates were only 5% with the use of triple csDMARD therapy, but more than three times as high with the use of a bDMARD at week 24.[13] This and other data informing the updated EULAR recommendations have been provided in our publication.[2]

Still, we are fully aware that an optimal therapeutic approach may not be affordable in certain countries. In these countries, deviations from the optimal route may have to come into place, as has been suggested by Asia Pacific League of Associations for Rheumatology (APLAR).[14] Of note, the current task force included representatives from Asia, Latin America and North America, including an APLAR president.

The comments presented and references provided by Misra *et al* raise a number of additional points, which have been quite clearly addressed in the recommendations. First, do Misra *et al* use rapid escalation of MTX to 25 mg (weight adjustment may have to be done)? This has been shown to lead to a state of low disease activity in 40% of patients even without GC use[15]; on addition of GC, the rate of low disease activity may be even much higher, as mentioned above. Second, do Misra *et al* follow their patients regularly using validated disease activity measures and do they record the results as well as the treatment target and the shared decision making with their patients in their charts? And third, do Misra *et al* indeed intensify therapy if at least 50% improvement is not attained at 3 months or the target not reached at 6 months? In line with the T2T strategy,[12] the last two points are of utmost importance, irrespective of the availability of specific drugs.[12] In this context, we would highly recommend to use short-term GCs on switching to or adding another csDMARD, in line with the recommendation provided for patients with lack of bad prognostic markers. Finally, it should be borne in mind that the availability of biosimilar (bs) DMARDs may increase the accessibility of bDMARD and tsDMARD therapies, since they may put pressure on current drug prices. Rheumatologists in countries like India also have a responsibility to point drug manufacturers to this source of inequity. The aspect of bsDMARDs has been addressed in the EULAR recommendations too.

Thus, overall, the EULAR recommendations are based on the evidence provided by three SLRs. The fact that representatives from all regions of the world were included in the task force ensured that regional aspects would be accounted for. However, the EULAR recommendations are 'recommendations' and may be used as a guidance to develop local recommendations in accordance with respective needs. That these may not be always fully in line with the core recommendations, and that some items may have to be adjusted to local needs and financial constraints, is well understood. However, the reasons for such deviation should be clearly stated and not be based on constructed evidence.

Josef S Smolen,[1] Robert B M Landewé,[2] Désirée van der Heijde[3,4]

[1]Division of Rheumatology, Department of Medicine 3, Medical University of Vienna, Vienna, Austria
[2]Rheumatology, Atrium Medical Center Heerlen, Heerlen, The Netherlands
[3]Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[4]Rheumatology, Diakonhjemmet Hospital, Oslo, Norway

Correspondence to Professor Josef S Smolen, Rheumatology, Medical University of Vienna, Vienna 1090, Austria; josef.smolen@wienkav.at

Competing interests None declared.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Smolen JS, Landewé RBM, van der Heijde D. *Ann Rheum Dis* 2017;**76**:e48.

Received 28 March 2017
Accepted 30 March 2017
Published Online First 11 April 2017

► http://dx.doi.org/10.1136/annrheumdis-2017-211446

Ann Rheum Dis 2017;**76**:e48. doi:10.1136/annrheumdis-2017-211455

REFERENCES

1 Misra D, Agarwal V, Sjarma A, *et al*. Update of the EULAR recommendations for the management of rheumatoid arthritis: a utopia beyond patients in developing countries? 2016.
2 Smolen JS, Landewé R, Bijlsma J, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017:annrheumdis-2016-210715.
3 Nam JL, Takase-Minegishi K, Ramiro S, *et al*. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2016 update of the EULAR recommendations for the management of rheumatoid arthritis. *Ann Rheum Dis* 2017:annrheumdis-2016-210713.
4 Ramiro S, Sepriano A, Chatzidionysiou K, *et al*. Safety of synthetic and biological DMARDs: a systematic literature review informing the 2016 update of the EULAR recommendations for management of rheumatoid arthritis. *Ann Rheum Dis* 2017:annrheumdis-2016-210708.
5 Chatzidionysiou K, Emamikia S, Nam J, *et al*. Efficacy of glucocorticoids, conventional and targeted synthetic disease-modifying antirheumatic drugs: a systematic literature review informing the 2016 update of the EULAR recommendations for the management of rheumatoid arthritis. *Ann Rheum Dis* 2017. doi: 10.1136/annrheumdis-2016-210711 [Epub ahead of print 29 Mar 2017].
6 Scott DL, Ibrahim F, Farewell V, *et al*. Tumour necrosis factor inhibitors versus combination intensive therapy with conventional disease modifying anti-rheumatic drugs in established rheumatoid arthritis: tacit non-inferiority randomised controlled trial. *Bmj* 2015;350:h1046.
7 Kiely P, Walsh D, Williams R, *et al*; Early Rheumatoid Arthritis Network. Outcome in rheumatoid arthritis patients with continued conventional therapy for moderate disease activity--the early RA network (ERAN). *Rheumatology* 2011;50:926–31.
8 Landewé R, Gorter S, Gaujoux-Viala C, *et al*. On publication policy, combination therapy, and the European league against rheumatism recommendations for the management of rheumatoid arthritis: comment on the article by Graudal et al. *Arthritis Rheum* 2011;63:3182–5.
9 Katchamart W, Trudeau J, Phumethum V, *et al*. Methotrexate monotherapy versus methotrexate combination therapy with non-biologic disease modifying anti-rheumatic drugs for rheumatoid arthritis. *Cochrane Database Syst Rev* 2010;4:CD008495.
10 Klarenbeek NB, Güler-Yüksel M, van der Kooij SM, *et al*. The impact of four dynamic, goal-steered treatment strategies on the 5-year outcomes of rheumatoid arthritis patients in the BeSt study. *Ann Rheum Dis* 2011;70:1039–46.
11 Verschueren P, De Cock D, Corluy L, *et al*. Methotrexate in combination with other DMARDs is not superior to methotrexate alone for remission induction with moderate-to-high-dose glucocorticoid bridging in early rheumatoid arthritis after 16 weeks of treatment: the CareRA trial. *Ann Rheum Dis* 2015;74:27–34.
12 Smolen JS, Breedveld FC, Burmester GR, *et al*. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;75:3–15.
13 O'Dell JR, Mikuls TR, Taylor TH, *et al*; CSP 551 RACAT Investigators. Therapies for active rheumatoid arthritis after methotrexate failure. *N Engl J Med* 2013;369:307–18.
14 Lau CS, Chia F, Harrison A, *et al*. APLAR rheumatoid arthritis treatment recommendations. *Int J Rheum Dis* 2015;18:685–713.
15 Emery P, Bingham CO, Burmester GR, *et al*. Certolizumab pegol in combination with dose-optimised methotrexate in DMARD-naïve patients with early, active rheumatoid arthritis with poor prognostic factors: 1-year results from C-EARLY, a randomised, double-blind, placebo-controlled phase III study. *Ann Rheum Dis* 2017;76.