EXTENDED REPORT

Development of EULAR recommendations for the reporting of clinical trial extension studies in rheumatology

Maya H Buch,[1,2] Lucia Silva-Fernandez,[3] Loreto Carmona,[4] Daniel Aletaha,[5]
Robin Christensen,[6] Bernard Combe,[7] Paul Emery,[1,2] Gianfranco Ferraccioli,[8]
Francis Guillemin,[9] Tore K Kvien,[10] Robert Landewe,[11] Karel Pavelka,[12]
Kenneth Saag,[13] Josef S Smolen,[5,14] Deborah Symmons,[15,16]
Désirée van der Heijde,[17] Joep Welling,[18] George Wells,[19] Rene Westhovens,[20,21]
Angela Zink,[22] Maarten Boers[23]

ABSTRACT

Objectives Our initiative aimed to produce recommendations on post-randomised controlled trial (RCT) trial extension studies (TES) reporting using European League Against Rheumatism (EULAR) standard operating procedures in order to achieve more meaningful output and standardisation of reports.

Methods We formed a task force of 22 participants comprising RCT experts, clinical epidemiologists and patient representatives. A two-stage Delphi survey was conducted to discuss the domains of evaluation of a TES and definitions. A '0–10' agreement scale assessed each domain and definition. The resulting set of recommendations was further refined and a final vote taken for task force acceptance.

Results Seven key domains and individual components were evaluated and led to agreed recommendations including definition of a TES (100% agreement), minimal data necessary (100% agreement), method of data analysis (agreement mean (SD) scores ranging between 7.9 (0.84) and 9.0 (2.16)) and reporting of results as well as ethical issues. Key recommendations included reporting of absolute numbers at each stage from the RCT to TES with reasons given for drop-out at each stage, and inclusion of a flowchart detailing change in numbers at each stage and focus (mean (SD) agreement 9.9 (0.36)). A final vote accepted the set of recommendations.

Conclusions This EULAR task force provides recommendations for implementation in future TES to ensure a standardised approach to reporting. Use of this document should provide the rheumatology community with a more accurate and meaningful output from future TES, enabling better understanding and more confident application in clinical practice towards improving patient outcomes.

INTRODUCTION

A randomised controlled trial (RCT) is the most objective means of evaluating an intervention and underpins regulatory decision-making and, if appropriate, the introduction of therapies into clinical practice. Many benefits of RCTs have been seen in the specialty of rheumatology, and particularly in the management of rheumatoid arthritis (RA).[1–10] While the aim of RCTs is to demonstrate the efficacy and safety of an experimental agent, their observation period typically spans a relatively short time frame. However, the use of therapies in chronic diseases necessitates more long-term evaluation. The introduction of new disease modifying anti-rheumatic drug (DMARD) therapies for the treatment of RA has been associated with a significant number of post-RCT extension studies,[11–17] henceforth termed 'trial extension studies' (TES), to report the longer-term outcomes of an experimental agent.

Role of TES

TES can evaluate in particular, the effects of cumulative exposure to a drug, capturing events through systematic reporting, monitoring of source data, and consistent coding, thus enabling further assessment of the long-term safety profile observed during the RCT.[18]

An additional benefit of TES that is cited is continued access to an effective but otherwise unlicensed treatment by RCT participants. However, since a favourable effect of the treatment may not have been clearly determined at the time of TES participation (with results from the preceding RCT and/or indeterminate prior studies not available), this raises legitimate ethical issues about the appropriateness of exposing patients to potentially ineffective or only partially effective treatments for additional periods of time.

Challenges of TES

While TES play a valid role, there are clear limitations that should be considered and potential weaknesses in the design and method of analysis that should be addressed.[19] TES benefit from the systematic reporting on cumulative drug exposure but have clear limitations in the detection of rare and unexpected events. In addition, selection bias associated with TES populations and lack of generalisability are key factors. These issues are discussed in more detail in the online supplementary material.

This makes interpretation challenging and sometimes unreliable. While guidance for reporting of RCTs[20 21] and safety data from biological DMARD registers[22] are available, no recommendations for TES in rheumatology have been published to date.[23] With this in mind, a task force was created with the principal aim of developing practical recommendations on key aspects of TES on the basis of the European League Against Rheumatism (EULAR) standard operating procedures,[24] and thereby a recommended standardised format for future TES data reporting to achieve greater transparency. This manuscript reports the final recommendations as agreed by the task force.

METHODS

The task force agreed that a systematic literature review was not indicated for this initiative, as it would merely serve to further establish the lack of consistency in TES and emphasise the need for the development of a standard for future application.[25]

The target population for these recommendations was chosen to be rheumatologists, trialists and researchers working in the field of rheumatology, patient organisations and policymakers. The general approach to this project followed the EULAR standardised operating procedures for the elaboration and implementation of evidence-based recommendations.[24]

The two task force conveners (MHB and MB) set up a multidisciplinary task force with participants selected based on their field of expertise, knowledge and experience as well as appropriate geographical distribution, primarily across Europe but also North America.

A first meeting of all task force members was convened in January 2011 to primarily define the domains for evaluation. This comprised two breakout sessions, with the task force split into two groups. Each group had a rapporteur who reported the outcome to the whole task force. After a final round of discussion, the task force agreed on the individual items for inclusion in a Delphi exercise. The Delphi method offers a consensus method that is widely used in health service research.[26] The two-step Delphi exercise for this initiative was web based, which permitted opinions to be provided and votes on the level of agreement to be cast independently and anonymously. Geographical limitations were also avoided by this approach. It was designed by LS-F and reviewed and modified as indicated by MHB, LC and MB. Details on how the Delphi exercise was formulated, responses were scored and the approach for informing final recommendations was devised can be found in the online supplementary material.

RESULTS

Task force composition

The multidisciplinary task force comprised 22 participants consisting of 17 rheumatologists, of whom six were clinical epidemiologists and 11 clinical trialists/expert clinicians, two biostatisticians, one fellow and two patient representatives. Participants represented 10 European countries, the USA and Canada.

Response rate

Of the 22 invited experts, three could not attend the first meeting (January 2011) but were subsequently apprised of the discussion and participated in the Delphi exercise. One of the patient representatives could not continue participation after the first meeting. Twenty of the 21 participants responded to the first and all 21 responded to the second Delphi exercise.

The two-step Delphi exercise was completed by January 2012, with subsequent analysis and dissemination of draft recommendations in March 2012. Final voting took place in May 2012. However, subsequent steps of involving additional stakeholders (see 'Results' section) and a meeting to discuss the recommendations (June 2013) led to a delay in establishing the recommendations for the purposes of submission. The task force approved this final document that included some modifications following the last step. More details on the timelines, responses and involvement of other stakeholders are detailed in the online supplementary material.

Domains for evaluation

At the initial meeting, the task force agreed on seven main domains to form the basis of the exercise. These are listed in box 1 with components within each domain that we wished to cover.

Final results

Percentage agreement for each recommendation (following the second Delphi exercise) is given. Where appropriate, mean (SD) scores have also been provided. Median (range) scores were also calculated and are included in the online supplementary material.

Definition of a TES

▶ *Study design definition* (100% agreement): A TES is a study that follows all patients beyond a pre-specified trial period whether the trial was (a) a placebo-controlled RCT with the possibility to cross over to an open-label experimental drug or (b) a placebo-controlled RCT with the possibility to cross over to usual care or (c) an active comparator RCT.

▶ *Start of a TES* (100% agreement): Should be stated in the pre-specified protocol with clear justification, and should be at the point of exposure to the experimental drug of interest. For the experimental randomised arm, this will be the start of the original RCT, while for those randomised to placebo/active comparator arm, this point will be on switching to experimental treatment.

Box 1 The key domains underpinning the Delphi exercise

1. Definition of a trial extension study (TES)
   Study design definition
   Definition of start of TES
   Duration of TES
   Patient population of TES
2. Development of a checklist of minimal data items/outcome necessary for a TES
   Minimal information a TES should collect
   Elements not amenable to accurate assessment by a TES
   Safety elements that may be elicited
   Efficacy
3. Additional data/outcomes
   Additional legitimate outputs from a TES
4. Method of analysis
5. Method of reporting results
   Inclusion of a flowchart
   Detail minimal standards by way of a checklist
   Frequency and nature of TES
6. Ethics and obtaining consent
7. Over-arching principles
   Consultation and stakeholder involvement
   General comments on TES and its reporting
   Sources of bias and generalisability

- *Minimum duration of a TES* (100% agreement): It was agreed by consensus not to define this; nevertheless, the rationale for the length chosen should be stated in the pre-defined protocol with adequate justification.
- *Population for inclusion in a TES* (96% agreement): Should include all patients included in the RCT, with the ability to separately report on patients who are of specific interest, for example, those in remission or with low disease activity.

Checklist of minimal data items/outcome necessary for a TES
Minimal information
The minimal information that should be collected and reported by a TES is listed in box 2. Minimum and maximum mean (SD) scores following the first Delphi exercise were 7.2 (2.72) and 9.9 (0.36) (refer to the online supplementary material for individual mean and median scores) with agreement by 100% of the task force in the second Delphi exercise.

The entire group also accepted the following statements relating to the nature of the initial RCT design following the first Delphi exercise:

Box 2 Minimal information to be included in a TES report

- Progress of subjects at each stage from RCT start to TES completion with:
  - A flow diagram detailing *absolute* numbers of subjects at each relevant time-point
  - Duration of active treatment
  - Time of last observation
- All drop-outs detailed
- The drop-out rates from each arm during the original RCT and the cross-over groups
- Reason for exclusion from the TES if the patient discontinues the drug
- Reason for cessation of follow-up
- Specification of reasons for cessation of follow-up other than adverse event or inefficacy as above, for example, geographical or doctor-related reasons
- Functional status at the time of inclusion in the TES if applicable
- Functional status at last observation if applicable
- Disease activity at the time of inclusion in the TES if applicable
- Disease activity at last observation if applicable
- For those patients entering the TES having achieved low disease activity or remission during the RCT, the sustainability of each disease state should be evaluated and reported
- For those subjects who enter a TES not having achieved remission/an acceptable disease activity state following the RCT, the number who achieve this during the TES should be reported to determine whether longer drug exposure has the potential to improve the disease state of such subjects further
- Disease-related co-medication (DMARD, corticosteroid) at each stage from RCT start to TES completion
- Any serious adverse events and outcome related to safety at each stage from RCT start to TES completion

DMARD, disease modifying anti-rheumatic drug; RCT, randomised controlled trial; TES, trial extension studies.

- The minimum data requirements for TES following placebo- and active comparator RCTs should be the same (93% agreement).
- A TES that follows an active comparator RCT should follow all randomised patients for the same period of time (not only patients on the experimental treatment) and including patients who may switch to an active comparator treatment (analysed separately) (mean (SD) 7.9 (2.23), median (range) 8.5).[2–10]

Safety and efficacy outcomes
Evaluation of safety aspects includes several elements, some of which it may not be feasible to capture within certain study designs. The following statements were agreed during the first Delphi exercise (minimum and maximum mean score of 7.0 and 8.4; refer to online supplementary material for individual scores) with 90% accepting all statements in the second round.
Safety
- TES may identify new adverse effects that the original RCT was not able to detect due to greater cumulative drug exposure.
- TES may identify whether the *incidence* of known adverse effects changes with longer-term drug exposure.
- TES may confirm whether the *nature* of known adverse effects identified from the RCT changes with longer-term exposure.
- TES are sub-optimal to detect rare safety events because they are not powered for this.
- TES are sub-optimal to detect rare safety events because they include a selected population (responders with likely no previous serious adverse events).
Efficacy
- Greater cumulative exposure to the active drug per patient in a TES might identify additional information on the drug's efficacy.
- While definitions of relapse are currently not available and require further work, if/when validated, a TES might allow evaluation of relapse including time to relapse and therefore the sustainability of original disease control.

Additional data/outcomes
- Economic evaluation of long-term treatment with the active drug may be possible if appropriate measures are recorded in the TES.
- A TES could not accurately evaluate health-related quality of life.

Method of analysis
Following the second Delphi exercise, this section required further iterations to refine the initial Delphi statements. These are detailed in box 3. Minimum and maximum scores of agreement were 7.3 and 9.4 (refer to the online supplementary material for individual scores).

Method of reporting results
Inclusion of a flowchart
- All TES reports should include a flowchart.
This was agreed as a minimal piece of information to accurately illustrate the treatment arms, and changes in treatment and in patient numbers during the course of the study (mean (SD) 9.9 (0.36)).
- In particular, the absolute measure/count should be reported (with/without the percentage).

Recommendation

Box 3 Guidance on data management and statistical approach statement

- The null hypothesis should be stated at the start where appropriate.
- Multiple comparisons should be taken into account when determining the level of statistical significance.
- The null hypothesis should take account of the results of the original RCT. Depending on the research question, the results of an RCT should be accommodated in the TES.
- The report should comment on cumulative outcome analysis (beneficial and adverse events) maintaining the original trial groups, that is, from RCT start not TES start, to avoid reporting of only the sub-selected patient group that proceeds to the TES.
- The selection bias associated with a TES population means meaningful non-inferiority/superiority analysis would not be reliable. The report should focus on how data for sustained effect from the start to the end of the TES period, within a single group or the difference between groups was analysed and whether there was any suggestion of increased effect (although this could not be subject to formal statistical testing).
- The plan for subjects that drop out of a TES should be specified to demonstrate sustained effect from the start to the end of the TES period. With reducing number of participants (the denominator), the proportion responding will artificially increase if/when the number of patients (numerator) responding stays the same.
- The analysis should include survival/retention rates on therapy explicitly reporting the number of patients at each milestone with reasons for change detailed.
- A plan on how to analyse this should be included with both intent-to-treat (ITT) (denominator as the original number entering the RCT) and completer (those entering TES only) population analyses reported. A completer analysis should always be reported together with an ITT analysis.
- The repeated measures analysis of the data from a TES in rheumatology should include the area under the curve of absolute disease activity (ie, not dichotomous response/change) preferentially expressed as a score (eg, DAS, SDAI, etc).
- A TES should preferably include hard endpoints (eg, death, work disability, joint replacement surgery, hospital admission) from the TES with or without linkages with other data sources.

RCT, randomised controlled trial; TES, trial extension studies. The agreement scores were recorded after Delphi round 1.

In a TES, the denominator of a cohort typically decreases over time, which results in the reporting of (artificial) increasing percentages of response rates over time.[27] The use of absolute numbers ensures accurate synthesis of the data.

Figure 1 includes a schematic of suggested flowcharts for either placebo-controlled or active comparator RCTs that was accepted by the group (mean (SD) 9.0 (2.06)).

Frequency and nature of reporting outputs from a TES
The following recommendations were made (mean scores between 8.2 and 8.8; refer to the online supplementary material for individual scores):

- Reporting frequency should not be specified for all TES since this depends on the research question.
- However, the protocol of each TES should pre-specify the minimum frequency of reports to be written and the basis for them (purpose, outcomes, length of RCT).
- The efficacy and safety results of a TES should generally be reported together; abstract selection committees and journal editors should carefully consider reporting of efficacy alone before acceptance.

Consent
The recommendations related to obtaining consent are detailed below; this item in particular required specific input from the patient representative (refer to the online supplementary material for individual scores on additional questions that had means scores of between 6.2–9.4).

- All of the subjects undergoing an RCT should be informed of the importance of long-term surveillance and be given the opportunity of entering in the long-term follow-up (mean (SD) 9.4 (0.85)).
- Subjects should sign a new consent form both for continuation of the drug and for data collection at that time point (mean (SD) 7.6 (2.87)).
- Annual updates for consent are not recommended (mean (SD) 3.7 (4.4)).

Over-arching principles
- The report of a TES should be consistent with and consolidate existing established guidelines including CONSORT[20][28] and STROBE[29] (mean (SD) score 9.4 (0.85)).
- The report of a TES should be consistent with the ACR/EULAR recommendations on the reporting of clinical trials in RA[21] (mean (SD) score 8.9 (1.88)).

General comments on TES and its reporting
All the following statements were accepted by 95% of the group in the second Delphi exercise, agreement with the individual statements having been established as part of the initial Delphi exercise (agreement score out of 10):

- While data linkage is important for long-term observation, access may be difficult as pharmaceutical companies conduct most TES; this may in turn limit the overall benefit of such studies (mean (SD) 7.1 (2.06)).
- TES, by definition, comprise a sub-selected population, not reflective of routine care; hence, even if all patients in an RCT were entered into a TES, such a study is generalisable only to patients with similar disease characteristics (mean (SD) 7.9 (1.76)).
- The absence of a clear null hypothesis may make the definition of comparator groups in a TES difficult (mean (SD) 7.4 (1.74)) and should therefore be stated where appropriate (see table 3 for details on method of data analysis).

Potential sources of bias or lack of generalisability
Several factors were identified as possibly influencing the inclusion of patients in a TES following completion of an RCT, which could introduce sources of bias and lack of generalisability (80% agreement to include all the following statements):

- The requirement of a certain level of response (mean (SD) 7.9 (2.67))
- The stage of the disease of the patient (mean (SD) 7 (2.18)).
- The fact that the investigator is remunerated for each patient recruited or that the patients may also receive financial

Figure 1 Schematic of flowcharts for inclusion in any trial extension study (TES) report following (A) a placebo-controlled randomised controlled trial (RCT) or (B) an active comparator trial. FU, follow-up; W, withdrawal. *And loss to follow-up number. The TES flowchart should detail the numbers from the start of the original RCT to the end of the TES. 'W' and the vertical dotted line shown between the RCT and TES phases denote those subjects who complete the RCT but opt not to proceed to the TES. During the RCT and TES, numbers of early withdrawals and patients subsequently lost to follow-up should also be included. (A) In the placebo-controlled RCT stage, subjects in the experimental arm who withdraw and proceed to standard of care, and in the placebo arm, subjects who withdraw and proceed to experimental or standard of care should be included. (B) Similarly, in the active comparator trial, subjects in the experimental arm who withdraw and proceed to standard of care, and in the active arm, subjects who withdraw and proceed to experimental or standard of care, should be included.

compensation and that the drug is free of charge could be of importance in some health systems (mean (SD) 7.4 (1.7)).
▶ Geographical differences in practice/approach (leading to differences in the number and nature of patients included) (mean (SD) 7.5 (2.45)).
▶ Unwanted heterogeneity from countries where treatment options may be more limited (eg, patients with higher levels of disease activity recruited where otherwise only patients in remission/with low disease activity would be included) (mean (SD) 7.6 (1.45)).

Consultation on recommendations and stakeholder involvement
The Delphi process established whether input from relevant stakeholder organisations, namely, industry, regulatory authorities (Food and Drug Administration (FDA), European Medicines Agency (EMA)) and contract research organisations (CRO) should be sought. In the initial Delphi exercise, 75% voted in favour of some level of industry input, 94% for regulatory authorities and 81% for CRO.

The second Delphi exercise asked for agreement that each of these organisations be included in the initiative:
▶ Industry and regulatory authority input into the final recommendations was recommended, with mean (SD) scores out of 10 of 7.2 (2.48), 8.3 (1.77) and 4.9 (2.85) recorded for the FDA, EMA and CRO, respectively.

Key industry companies that have been associated with new drugs in the RA arena were therefore approached (refer to online supplementary material for details of the companies represented).

DISCUSSION
We present a series of pragmatic recommendations on the design and reporting of TES in rheumatological conditions (mainly inflammatory arthritis, although the basic principles are generally applicable), based on a high degree of expert consensus. Our EULAR task force comprised a group of experts encompassing a range of expertise including clinical trialists, clinicians experienced in RA treatment, and clinical epidemiologists as well as patient representatives. A wide range of countries

and health systems were represented, albeit with some omissions (eg, absence of individuals from Asia), although the opportunity to evaluate these recommendations in the wider community in the future should highlight any differing perspectives. With a generally accepted methodology for prospective observational studies, we felt an additional systematic review was not necessary and decided to use our expert opinion to formulate guidance for TES. These recommendations complement those established for clinical trials[21] and registries.[22]

Central to the recommendations was the principle that a TES report should focus on cumulative outcome analysis, maintaining the original trial groups to avoid reporting of only the subselected patient group that proceeds to the TES, and thereby achieve better generalisability of results. Furthermore, the task force was clear that absolute numbers and not just percentage response rates should be reported. To facilitate this, we recommend a flow diagram detailing absolute numbers of subjects at each relevant time point, with clear illustration of drop-outs and the reason for cessation and/or exclusion at each relevant stage. While it was agreed that a TES might elaborate on the incidence and nature of adverse events over time, they are not designed to capture rare safety signals. TES reports may also have the potential to inform on the durability of response and the dynamics of achieving pre-determined targets of treatment (low disease activity and remission). It was agreed that any analysis should be pre-specified in the protocol but should always include an intention-to-treat in addition to a completer approach. We acknowledge there are elements that may in particular be the subject of further discussion in the wider community, for example, the issue of split reporting. While the task force discouraged this, each case should be considered individually as there may be instances when there is utility in this approach to ensure relevant data that is of interest is disseminated within the public domain.

The recommendations were actively commented on by several industry companies (see the 'Consultation of recommendations and stakeholder involvement' section) and include their specific feedback (which has been indicated directly in the results where appropriate in the online supplementary material) and as such, gained the approval of the participating stakeholders. While EMA representation did not suggest changes to the recommendations, it acknowledged the importance of standardisation. The interaction also highlighted how regulatory expectations may drive the industry approach on whether and how TES should be undertaken.

While we acknowledge that the working group was perhaps relatively small for a consensus exercise, following dissemination of these recommendations, we would anticipate a subsequent exercise to capture how they have been received in the wider rheumatology, trial and industry communities. In future, it will be important for journal reviewers and editors to measure future TES reports against the standard set by these recommendations. The future research agenda will include a systematic review of forthcoming TES to evaluate how well this document is utilised, with further refinement based on the nature of outcomes observed. In addition, regulatory agencies may wish to consider the recommendations and associated issues and how these may influence their expectations from industry. This initiative and the interactive session at EULAR, Madrid 2013 with relevant stakeholders will hopefully be a springboard for further action (the outcome of the EULAR meeting is summarised in the online supplementary material).

In summary, there is a clear unmet need for a reliable approach to the reporting of TES to maximise our understanding of drug effects in chronic conditions. This initiative, its principles and resulting recommendations apply to TES for any drug in RA as well as for drugs used to treat other chronic rheumatological conditions. This document provides much needed first recommendations to ensure a transparent and standardised approach to the reporting of future TES.

Author affiliations
[1]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, UK
[2]NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[3]Rheumatology Department, Hospital Universitario de Guadalajara, Guadalajara, Spain
[4]Institute for Musculoskeletal Health, Madrid, Spain
[5]Department of Internal Medicine 3, Division of Rheumatology, Medical University of Vienna, Vienna, Austria
[6]Department of Rheumatology, Musculoskeletal Statistics Unit, The Parker Institute, Copenhagen University Hospital, Bispebjerg and Frederiksberg, Copenhagen, Denmark
[7]Department of Rheumatology, Lapeyronie Hospital, Montpellier I University of Montpellier, France
[8]Division of Rheumatology, Catholic University of the Sacred Heart, Rome, Italy
[9]Université de Lorraine, Université Paris Descartes, Nancy, France
[10]Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
[11]Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands
[12]Institute of Rheumatology and Clinic of Rheumatology, Charles University, Prague, Czech Republic
[13]Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham School of Medicine, Birmingham, Alabama, USA
[14]2nd Department of Medicine, Center for Rheumatic Diseases, Hietzing Hospital, Vienna, Austria
[15]Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK
[16]NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Trust, Manchester, UK
[17]Leiden University Medical Center, Department of Rheumatology, Leiden, The Netherlands
[18]EULAR Patient Research Partner, Ede, The Netherlands
[19]Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada
[20]Department of Development and Regeneration, KU Leuven, Skeletal Biology and Engineering Research Center, Leuven, Belgium
[21]Department of Rheumatology, University Hospitals Leuven, Leuven, Belgium
[22]German Rheumatism Research Centre, Berlin, Germany
[23]Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands

REFERENCES
1 O'Dell JR, Leff R, Paulsen G, et al. Treatment of rheumatoid arthritis with methotrexate and hydroxychloroquine, methotrexate and sulfasalazine, or a combination of the three medications: results of a two-year, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2002;46:1164–70.
2 Boers M, Verhoeven AC, Markusse HM, et al. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. *Lancet* 1997;350:309–18.

3 Maini R, St Clair EW, Breedveld F, et al. Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. ATTRACT Study Group. Lancet 1999;354:1932–9.

4 Weinblatt ME, Kremer JM, Bankhurst AD, et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. N Engl J Med 1999;340:253–9.

5 Weinblatt ME, Keystone EC, Furst DE, et al. Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. Arthritis Rheum 2003;48:35–45.

6 Cohen SB, Emery P, Greenwald MW, et al. Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. Arthritis Rheum 2006;54:2793–806.

7 Genovese MC, Becker JC, Schiff M, et al. Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. N Engl J Med 2005;353:1114–23.

8 Emery P, Keystone E, Tony HP, et al. IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial. Ann Rheum Dis 2008;67:1516–23.

9 Smolen J, Landewe RB, Mease P, et al. Efficacy and safety of certolizumab pegol plus methotrexate in active rheumatoid arthritis: the RAPID 2 study. A randomised controlled trial. Ann Rheum Dis 2009;68:797–804.

10 Keystone EC, Genovese MC, Klareskog L, et al. Golimumab, a human antibody to tumour necrosis factor {alpha} given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: the GO-FORWARD Study. Ann Rheum Dis 2009;68:789–96.

11 Nuki G, Bresnihan B, Bear MB, et al. Long-term safety and maintenance of clinical improvement following treatment with anakinra (recombinant human interleukin-1 receptor antagonist) in patients with rheumatoid arthritis: extension phase of a randomized, double-blind, placebo-controlled trial. Arthritis Rheum 2002;46:2838–46.

12 Nishimoto N, Miyasaka N, Yamamoto K, et al. Long-term safety and efficacy of tocilizumab, an anti-IL-6 receptor monoclonal antibody, in monotherapy, in patients with rheumatoid arthritis (the STREAM study): evidence of safety and efficacy in a 5-year extension study. Ann Rheum Dis 2009;68:1580–4.

13 Keystone E, Fleischmann R, Emery P, et al. Safety and efficacy of additional courses of rituximab in patients with active rheumatoid arthritis: an open-label extension analysis. Arthritis Rheum 2007;56:3896–908.

14 Genovese MC, Bathon JM, Fleischmann RM, et al. Longterm safety, efficacy, and radiographic outcome with etanercept treatment in patients with early rheumatoid arthritis. J Rheumatol 2005;32:1232–42.

15 Westhovens R, Kremer JM, Moreland LW, et al. Safety and efficacy of the selective costimulation modulator abatacept in patients with rheumatoid arthritis receiving background methotrexate: a 5-year extended phase IIB study. J Rheumatol 2009;36:736–42.

16 Scott DL, Smolen JS, Kalden JR, et al. Treatment of active rheumatoid arthritis with leflunomide: two year follow up of a double blind, placebo controlled trial versus sulfasalazine. Ann Rheum Dis 2001;60:913–23.

17 Wollenhaupt J, Silverfeld J, Bong Lee E, et al. Tofacitinib, an oral Janus kinase inhibitor, in the treatment of rheumatoid arthritis: open-label, long-term extension safety and efficacy up to 5 years. Arthritis Rheum 2013;65(S2328).

18 Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401–6.

19 Buch MH, Aletaha D, Emery P, et al. Reporting of long-term extension studies: lack of consistency calls for consensus. Ann Rheum Dis 2011;70:886–90.

20 Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg 2012;10:28–55.

21 Aletaha D, Landewe R, Karonitsch T, et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Ann Rheum Dis 2008;67:1360–4.

22 Dixon WG, Carmona L, Finckh A, et al. EULAR points to consider when establishing, analysing and reporting safety data of biologics registers in rheumatology. Ann Rheum Dis 2010;69:1596–602.

23 Christensen R, Bliddal H, Henriksen M. Enhancing the reporting and transparency of rheumatology research: a guide to reporting guidelines. Arthritis Res Ther 2013;15:109.

24 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.

25 Vandenbroucke JP. STREGA, STROBE, STARD, SQUIRE, MOOSE, PRISMA, GNOSIS, TREND, ORION, COREQ, QUORUM, REMARK... and CONSORT: for whom does the guideline toll? J Clin Epidemiol 2009;62:594–6.

26 Jones J, Hunter D. Consensus methods for medical and health services research. BMJ 1995;311:376–80.

27 Kristensen LE, Saxne T, Geborek P. The LUNDEX, a new index of drug efficacy in clinical practice: results of a five-year observational study of treatment with infliximab and etanercept among rheumatoid arthritis patients in southern Sweden. Arthritis Rheum 2006;54:600–6.

28 Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008;337:a2390.

29 von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007;335:806–8.