BMJ Open

Interpreting trial results following use of different intention-to-treat approaches for preventing attrition bias: A meta-epidemiological study protocol

SCHOLARONE™
Manuscripts

Interpreting trial results following use of different intention-to-treat approaches for preventing attrition bias: A meta-epidemiological study protocol

Anna Dossing[1,2], Simon Tarp[1], Daniel E. Furst[3], Christian Gluud[4], Joseph Beyene[5], Bjarke B. Hansen[1], Henning Bliddal[1], Robin Christensen[1]

Affiliations
1: Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospitals, Bispebjerg and Frederiksberg, Denmark
2: Department of Clinical Medicine, Faculty of Medical and Health Sciences, University of Copenhagen, Denmark
3: University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, USA
4: Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
5: Department of Clinical Epidemiology & Biostatistics, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

Correspondence to:
Dr. Robin Christensen, MSc, PhD
Head of Musculoskeletal Statistics Unit;
The Parker Institute, Department of Rheumatology.
Copenhagen University Hospital, Bispebjerg and Frederiksberg.
Nordre Fasanvej 57
DK-2000 Copenhagen Frederiksberg
Denmark
Phone: +45 3816 4165
Fax: +45 3816 4159
E-Mail: Robin.Christensen@frh.regionh.dk

Word count: (excluding title page, abstract, references, figures and tables): 4,654

ABSTRACT

Introduction: When subjects drop out of randomised clinical trials, as frequently happens, the intention-to-treat principle do not apply, potentially leading to attrition bias. Data lost from patient dropout/lack of follow-up are statistically addressed by imputing, a procedure prone to bias. Deviations from the original definition of ITT have led to various terminologies, such as modified ITT (mITT). As yet, the impact of the potential bias associated with mITT has not been assessed. Our objective is to investigate potential bias and disadvantages of performing mITT and evaluate possible concerns when executing different mITT approaches in meta-analyses.

Methods and analysis: Using meta-epidemiology on randomised trials considered less prone to bias (i.e., good internal validity) and assessing biological or targeted agents in patients with rheumatoid arthritis (RA), we will meta-analyse data from 10 biological and targeted drugs based on collections of trials that would correspond to 10 individual meta-analyses.

Ethics and dissemination: This study will enhance transparency for evaluating mITT treatment effects described in meta-analyses. The intended audience will include health care researchers, policymakers and clinicians. Results of the study will be disseminated by peer-review publication.

Protocol registration: In prospero CRD42013006702, 11. December 2013

2

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This meta-epidemiological study is the first to focus on bias associated with mITT and its impact on effect size.

- This study will influence how results from RCTs and meta-analysis should be interpreted; this study will provide a framework for designing, conducting and reporting future RCTs in order to minimise attrition bias.

- In practice, various definitions of both ITT and mITT are used. Some may argue that our definitions are not stringent enough or inverse overly stringent.

INTRODUCTION

Meta-epidemiological research

Inadequate quality of trials may distort the results from meta-analyses and systematic reviews [1]. Consequently, meta-epidemiological studies are carried out to quantitatively evaluate bias across many randomised clinical trials (RCTs) in different meta-analyses [2-4]. Overall flaws in the design, conduct, analysis, and reporting of RCTs can introduce bias—systematic errors that lead to overestimating or underestimating the benefits and the harms of treatment [5;6].

Methodological quality and risk of bias assessment

The Cochrane Collaboration's Handbook for Systematic Reviews of Interventions provides guidance to authors to critically review trial outcome using the risk of bias (RoB) assessment tool [5]. The RoB tool requires authors to evaluate the well-established limitations of RCTs, including sequence generation, allocation concealment, blinding of participants, personnel and outcome assessors, loss to follow-up with failure to apply the intention-to-treat (ITT) principle, and selective outcome reporting [5;7]. In the course of meta-epidemiological studies, other sources of bias in RCTs have been identified, such as significant discrepancies favouring intervention in single (rather than multicentre) trials, in trials with small (rather than large) sample sizes, and in using subjective (rather than objective) outcome measures [3;8-11]. Most recently, funding source has become a distinct possibility as a source of bias, with for-profit organisation funding likely favouring pro-intervention results [12-14].

The intention-to-treat principle

The ITT principle has two main rules: 1) it requires all participants in a RCT to be analysed according to their original allocation, regardless of their adherence to the trial protocol; and 2) all randomised participants must be included in the analysis. The first rule acts to conserve randomisation, which is executed to avoid selection bias and thereby produces treatment groups in which the distributions of prognostic factors—both known and unknown—are similar. The second rule serves to avoid attrition bias when evaluating a treatment assignment [15]. Attrition bias is attributed to systematic differences between groups in withdrawals from the study [7].

Frequently, not all participants in RCTs are analysed as they were initially randomised [16-19]. Various deviations from the definition of ITT have led to the vaguely defined term "modified ITT" (mITT), which may further compromise true randomisation. The incidence of trials reporting the use of mITT has increased over the years [20]. Trials that report mITT exclude patients from analysis post-randomisation compared to trials reporting on the ITT population [12]. Although post-randomisation exclusion is known to produce bias [21;22], the potential magnitude and direction of bias associated with mITT is unknown.

Based on the most common deviations from ITT described in the literature [17;18;20], mITT can be divided into four categories, as illustrated in **Figure 1**.
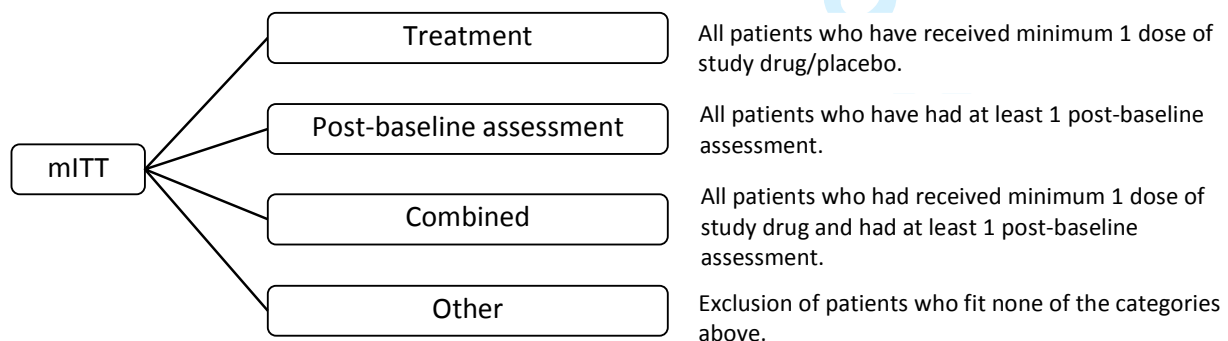

Figure 1: Overview of modified intention-to-treat (mITT) categories. The four categories are based on the most common deviations described in the literature.

Handling of missing data

Executing ITT and mITT analyses can be difficult, as missing data is common [23;24]. Missing data comprises single missing data points as well as missing datasets due to withdrawal [24]. Data can be missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). MNAR occurs when missing data depend on unobserved values: and the missing data may lead to bias [23-25]. Regardless of trial design and objective, it is likely that some of the data will be MNAR to a degree.

Problems from missing data in RCTs can be somewhat mitigated by data imputation—a procedure by which missing data are replaced with a presumably conservative estimate [25;26]. Two single imputation (SI) methods often used are "*last observation carried forward*" (LOCF) or "*baseline observation carried forward*" (BOCF). LOCF uses the last observed value in place of the missing outcome, whereas BOCF uses the baseline observation as the value replacing the outcome [25;27]. LOCF and BOCF have been widely criticised, as they are based on the assumption that lack of treatment equates with a halt in disease progression [19;28;29] and the use of SI risk introducing bias [30]. A newer and more promising imputation method is multiple imputation (MI), which involves creating several different plausible imputed datasets based on a Bayesian approach and then combining the results obtained from each of them [25;30;31].

Loss to follow-up is frequent in RCTs and can be attributed to a variety of causes [21;23]. Most patients drop out of trials due to lack of efficacy, adverse effects of treatment, or both [24]. When handling a dichotomous outcome in a responder analysis, it is often considered suitable to attribute patients' withdrawal to lack of efficacy, and therefore assume treatment failure [25;32] (also referred to as non-responder imputation (NRI)). NRI is applied inconsistently to different withdrawal populations, e.g. all patients who withdraw are considered treatment failures as opposed to only patients who withdraw due to lack of efficacy are considered treatment failures. In principle there should be no objection to applying NRI to different withdrawal populations [25] and no empirical evidence have documented whether it affects trial results.

Rationale for this meta-epidemiological study

Our extensive search did not find any previous systematic assessment of empirical evidence for bias associated with mITT [21;22]. Furthermore, mITT has not been assessed as to whether the type or number of modifications applied affects the estimated efficacy outcome differently. To investigate potential bias, we focus on trials assessing biological (or targeted) interventions in patients with rheumatoid arthritis (RA), as these trials are relatively recent and we anticipate that these yield reasonable internal validity. With regards to imputation, we wish to examine how imputation of missing data affects RA trials investigating clinical response, as this have, to our knowledge, not previously been done. The study also aims to shed a light on the effect of selectively applying NRI to different withdrawal populations, which so far has remained unexplored.

Biological and targeted agents for rheumatoid arthritis

RA is a chronic inflammatory autoimmune disease characterised by joint swelling and joint tenderness with destruction of synovial joints primarily affecting the hands [33-36]. The inflammatory load drives the destructive progression of the disease and leads to severe disability and premature mortality [37;38]. Diagnosis of RA requires that patients have a minimum of four criteria that persist for 6 months [33]. The seven qualifying criteria are: 1) morning stiffness; 2) arthritis of three or more joints or joint areas; 3) arthritis of hand joints; 4) symmetric arthritis; 5) rheumatoid nodules; 6) serum rheumatoid factor (RF); and 7) radiographic changes.

Treatment of RA encompasses multiple interventions. Though RA has been shown to progress over time despite treatment [37], early therapeutic intervention improves clinical outcomes and reduces the accrual of joint damage and disability [39]. RA can be managed using disease-modifying anti-rheumatic drugs (DMARDs). DMARDs form two major classes: synthetic chemical compounds (sDMARDs) and biological agents (bDMARDs). sDMARDs comprise the conventional DMARDs (csDMARDs) (e.g., methotrexat [MTX] and sulfasalazin) and the new targeted sDMARDs (tsDMARDs); e.g., tofacitinib. bDMARDs is a heterogeneous group of pharmaceuticals including abatacept, adalimumab, anakinra certolizumab, etanercept, golimumab, infliximab, rituximab, and

tocilizumab that help to control the autoimmune inflammation associated with arthritis [40-43]. The DMARDs of interest are summarised in **Table 1** [43].

DMARDs	Pharmaceutical
csDMARDs	methotrexat, sulfosalazin, leflunomid
tsDMARD	Tofacitinib
bDMARDs	adalimumab, abatacept, anakinra, certolizumab pegol, etanercept, golimumab, infliximab, rituximab, tocilizumab

Table 1: *Overview of disease-modifying anti-rheumatic drugs (DMARD) groups: conventional DMARDs (csDMARDs); the new targeted sDMARDs (tsDMARDs); and biological agents (bDMARDs).*

In RA RCTs on biological and targeted agents, the control group is commonly treated with MTX, as lack of treatment can lead to irreversible loss of physical function in RA patients [44;45]. Rescue therapy (e.g., regulation of dose or addition of MTX or bDMARD or tsDMARD) is acknowledged by both the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA) and is typically offered when treatment response is evaluated as inadequate [46;47]. Because patients who receive rescue therapy are commonly encountered as withdrawal patients in the trials protocols; rescue therapy challenges the ITT principle and may contribute to attrition bias.

The treatment effect in RA RCTs is typically evaluated using the American College of Rheumatology (ACR) response criteria and/or variations of the European Disease Activity Score (DAS28) [44]. The DAS28 score discriminates between high and low disease activity [34]. The ACR response criteria's definition of improvement in RA trials includes improvement in the joint counts and improvement in at least three of the five following: 1) patient assessment; 2) physician assessment; 3) erythrocyte sedimentation rate; 4) pain scale; and 5) functional questionnaire [48]. The EMA, the European League Against Rheumatism (EULAR), and the ACR consent that validated composite clinical outcomes such as the ACR response criteria should be used to document efficacy of treatment. Specifically the ACR20, ACR50, or ACR70 should be used to document signs and symptoms after 3–6 months, referring respectively to a 20%, 50%, or 70% improvement in disease activity [49-51].

Objective

Our primary objective is to examine

1. Whether mITT is associated with different effect sizes compared to ITT.

Secondarily we wish to examine

2a. How the choice of imputation technique influences the effect size.

2b. How different ITT populations affect effect size.

Hypotheses

We hypothesise that different ITT populations are associated with differences in treatment effect sizes. We expect the level of statistical significance will depend on the type of modification, the number of modifications applied, and the percentage of patients excluded from final efficacy analysis. Furthermore, we hypothesise that the use of different mITT approaches is strongly associated with industry funding [12] and these modifications are applied to artificially increase the effect size for the apparent benefit in trials directly related to a specific pharmaceutical company.

METHODS AND ANALYSIS

Protocol and registration

Our protocol is registered on PROSPERO (CRD42013006702)[52]; our manuscript conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for reporting systematic reviews and meta-analyses [53].

Eligibility criteria

All RA RCTs concerning EMA- or FDA-approved biologics and targeted agents evaluating efficacy as ACR20, ACR50, or ACR70 will be considered eligible, independent of whether they include ITT or mITT analysis. We will include published RCTs. Open-label studies will be excluded from our analysis because performance and detection bias are inherent when the outcome is not considered an objective endpoint [3].

Search and selection of trials and meta-analyses

We will search PubMed, EMBASE, The Cochrane Central Register of Controlled Trials (CENTRAL), and LILACS using a combination of keywords and text words related to RA. The PubMed search strategy is listed below. See **Appendix 1** for the EMBASE, the Cochrane Library, and LILACS search strategies. The World Health Organisation (WHO) Clinical trials Portal (ICTRP), clinicaltrials.gov, FDA, EMA and pharmaceutical companies' trial result databases will be searched to identify unpublished data.

PubMed search strategy: (((("Receptors, Tumor Necrosis Factor"[nm] OR TNFR:Fc OR "TNFR-Fc fusion protein"[Supplementary Concept] OR "TNFR-Fc fusion protein"[All Fields] OR "etanercept"[All Fields] OR "enbrel"[All Fields]) OR ("infliximab"[Supplementary Concept] OR "infliximab"[All Fields] OR "remicade"[All Fields] OR "mab ca2"[All Fields] OR "monoclonal antibody ca2"[All Fields]) OR ("adalimumab"[Supplementary Concept] OR "adalimumab"[All Fields] OR "humira"[All Fields]) OR ("interleukin 1 receptor antagonist protein"[MeSH Terms] OR "interleukin 1 receptor antagonist protein"[All Fields] OR "anakinra"[All Fields] OR "kineret"[All Fields] OR "antril"[All Fields]) OR ("abatacept"[Supplementary Concept] OR "abatacept"[All Fields]) OR CTLA4Ig[All Fields] OR "orencia"[All Fields]) OR ("rituximab"[Supplementary Concept] OR "rituximab"[All Fields] OR "rituxan"[All Fields] OR "idec c2b8"[All Fields]) OR ("golimumab"[All Fields] OR "golimumab"[Supplementary Concept] OR "simponi"[All Fields] OR "cnto-148"[All Fields] OR ("cnto"[All Fields] AND "148"[All Fields])) OR ("tocilizumab"[All Fields] OR "tocilizumab"[Supplementary Concept] OR "atlizumab"[All Fields] OR "actemra"[All Fields]) OR ("certolizumab"[All Fields] OR "certolizumab pegol"[Supplementary Concept] OR "CDP870"[All Fields] OR ("cdp"[All Fields] AND "870"[All Fields]) OR "cimzia"[All Fields]) OR ("tofacitinib"[Supplementary Concept] OR "tofacitinib"[All Fields]) OR ("Antibodies, Monoclonal"[Mesh] OR "Monokines"[Mesh] OR "Receptors, Interleukin-1"[Mesh] OR "Receptors, Interleukin-6"[Mesh])) AND ("Randomized Controlled Trial"[ptyp] OR "Controlled Clinical Trial"[ptyp] OR "Multicenter Study"[ptyp] OR "randomized"[tiab] OR "randomised"[tiab] OR "placebo"[tiab] OR "randomly"[tiab] OR "trial"[tiab] OR randomized controlled trials[mh] OR random allocation[mh] OR double-blind method[mh] OR single-blind method[mh]) AND ("Arthritis,

BMJ Open

Page 10 of 34

Rheumatoid"[MeSH Terms] OR (Rheumatoid[text word] AND arthriti*[text word]))) NOT (animals[mh] NOT human[mh]).

Date for final database update will be stated.

Data extraction

All RCT publications will be assigned an ID, and we will extract information on author, year of publication, journal of publication, company study name, company study number, and registration number. All characteristics will be typed into a custom made database (RHEUMATRIALS). We will extract baseline characteristics and inclusion criteria. RA will be classified on the basis of mean disease duration as early (≤ 6 months), established (6 months to 2 years), or late (> 2 years). Study duration until registration of primary outcome and duration of the longest placebo-controlled period will be extracted. We will extract type of primary outcomes (i.e., safety, efficacy, or both) and the name of the primary outcome (e.g., ACR20).

Trials will be classified based on the included patient populations. We will distinguish among three possibilities: whether patients have exhausted all csDMARD treatment options, whether they have had inadequate response to csDMARD treatment and are candidates for biologics or a targeted agent, or whether they have experienced an inadequate response to a bDMARD or tsDMARD.

Regarding the intervention, we will extract the main drug of interest, the allocation, the form of administration, dosing of the drug, and registration of whether dose is equivalent to standard dose. We will register if treatment with csDMARD concomitant to the intervention was allowed. Offering of rescue will be registered as "yes" or "no": time-point and the content of rescue will also be registered.

We will extract the total number of patients reaching the primary outcome and the total number of patients in the analysis. For efficacy outcomes, we will assess ACR20, ACR50, and ACR70. When available, we will extract the continuous efficacy outcome DAS28 as the mean change from baseline in each intervention group with the corresponding dispersion. If DAS28 was reported

based on C-reactive protein (CRP) and on erythrocyt sedimentation rate (ESR), we will extract only DAS28 (ESR), as both outcomes are in aggreement [54]. For safety outcomes, we will assess overall withdrawals, withdrawals due to adverse events, and number of serious adverse events.

Intention-to-treat analysis on efficacy outcome

We will extract the number of patients randomised in the trial in total, the number of randomised patients per intervention group, the number of patients completing each intervention group, and the number of patients included in efficacy analysis per intervention group. The method of efficacy analysis will be extracted on two levels. The first level will assess the *reported* protocolised method of analysis. The second level will register the *applied* method of analysis. The analysis population will be categorised as: "ITT," "mITT," "as observed," "per protocol," "other," or "unclear." A full ITT analysis will comprise both patients with clinically assessed outcome data (i.e., analyses conducted on all patients who completed the study) and patients with imputed outcome data (i.e., imputed data from all withdrawal patients). The specifics of each assessment are presented in **Table 2** [20].

Analysis population	Definition
ITT	All randomised patients are included in efficacy analysis and adverse outcome analysis.
mITT	All randomised patients, except a defined patient group, are included in efficacy analysis and adverse outcome analysis.
As observed (AO)	Only patients who complete the trial are included in efficacy analysis and adverse outcome analysis.
Per protocol (PP)	Patients who adhered to terms of eligibility, interventions, and outcome assessment pre-specified in the protocol are included in efficacy analysis and adverse outcome analysis.
Other	None of the above mentioned categories fits the analysis population.
Unclear	It is unclear which analysis is applied.

Table 2: *Analysis population for efficacy analysis and adverse outcome analysis* [12;17;20].

When the number of patients who complete the follow-up examination is the same as those originally randomised, we will register the applied method of analysis as ITT, regardless of intended protocolised analysis population. Data analyses will assess both the *reported* and the *applied* method of analysis.

Trials conducting 'as observed' or 'per protocol' analysis will be excluded from subsequent analysis stages (8a-9b, figure 2) as they do not assess missing data or withdrawal patients, *per se*.

Missing data will be assessed from two perspectives. We will extract how the trials handle single sets of missing data noted as "BOCF,", "LOCF," "combined," "other," or "unclear." "Combined" will refer to a combination of BOCF and LOCF, whereas "other" will refer to the use of other imputation techniques such as MI.

Subsequently we will determine whether trials distinguish among various withdrawal patients when handling missing data. We will register whether NRI (treatment fail) was applied to: a) patients who withdrew due to lack of efficacy; b) all patients who withdrew; or c) another defined patient group that withdrew but fit neither a or b.

When mITT is applied, we will assess the method of modification based on prior described categories (see **Figure 1** [12;17;20]). The modification will be registered as "Treatment" if final analysis comprised all randomised patients who had received at minimum one dose of the study drug. "Post-baseline assessment" will be noted as modification if final analysis included all patients who had at least one post-baseline assessment. The modification will be registered as "combined" if final analysis comprised all patients who had received at minimum one dose of the study drug and had at least one post-baseline assessment. Modifications that fail to fit into the three afore mentioned categories will be registered as "other." The number of modifications applied will be extracted as a numeric value.

A complete data extraction flowchart is presented in **Figure 2**.

Figure 2: *Data extraction flowchart.*

Evaluating methodological quality

The Cochrane risk of bias tool

The risk of bias within each trial will be assessed using the RoB tool as recommended by The Cochrane Collaboration (see below) [5;7]. Each domain will be rated as "low," "high," or "unclear" RoB. **Table 3** provides an overview of the Cochrane RoB components and their assessments, while **Table 4** provides an overview of other RoB components and their assessments. A domain will be rated as "unclear" if it fails to meet the criteria for "high" or "low" RoB.

To facilitate interpretation on the basis of the Cochrane risk of bias tool, each trial will subsequently be assigned an overall RoB. Overall RoB will be assessed tripartite as low risk (low for all key domains), high risk (high for ≥1 key domains), and unclear risk (unclear for ≥1 key domains) [9]. Overall RoB will also be assessed bipartite, categorized as low risk (low for all key domains) or high risk (high or unclear for ≥1 key domains) [3;55].

RoB item	Low RoB	High RoB
Sequence generation	It will be considered adequate if a random approach in the sequence generation process referred to a random number table, a random computer-generated number, coin tossing, drawing of lots, shuffling of cards, or throwing of dice. Multicentre trials described as randomised will be considered to have adequate sequence generation.	Date of birth, date of inclusion or admission, or record number of clinic/hospital is considered inadequate.
Allocation concealment	It will be considered adequate if there were no reasons to expect that the investigators responsible for inclusion were able to suspect which treatment was next. Both sequentially numbered, sealed, opaque envelopes and a central randomisation are considered adequate.	It will be regarded as inadequate if there is reason to expect that the investigators were able to suspect which treatment was next.
Blinding of patients, personnel, and outcome assessors	It will be considered adequate if the trials describe double-blinding.	It will be considered as inadequate if no blinding is described.
Incomplete outcome data	It will be considered adequate if missing data are few and distributed equally between intervention and control group. Further outcome data will be deemed adequate if data have been imputed using an appropriate technique and ITT analysis was applied.	It will be considered inadequate if it is unclear how many patients are included in final analyses. Further, it is considered inadequate if no imputation technique is applied or if it is unclear how extensive the missing dataset is (i.e., unclear how many patients withdrew).
Selective reporting	It will be considered adequate if the chosen efficacy outcome (ACR20, ACR50 and/or ACR70) is reported in accordance with the usual contemporary RA protocols and reported at all specified time-points if more than one time point exists.	It will be considered inadequate if the chosen efficacy outcome (ACR20, ACR50 and/or ACR70) is not reported in accordance with the usual contemporary RA protocols, or is not reported at all specified time-points if more than one time point exists.

Table 3: The Cochrane Risk of Bias tool.

Other risk of bias components

Funding will be registered according to funding source, as described in **Figure 3** [12]. Funding includes provision of manpower (authorship, statistical analysis, or other assistance), study materials (drug, placebo, assay kits, or similar materials), or grants [13]. For-profit organisations will be defined as companies that might acquire financial gain or loss depending on the outcome of the trial. Trials partly financed by for-profit agencies will be registered as co-financed.

As an extension to the funding aspect, we will assess whether conflict of interest is reported as "none," denoted by "Yes" or "No."

```
                        ┌──────────┐
                        │ Funding  │
                        └──────────┘
     ┌────────────┬───────────┼───────────┬─────────────┐
┌──────────┐ ┌────────────┐ ┌──────────────┐ ┌────────────┐ ┌────────────┐
│No funding│ │Not-for-profit│ │ Not reported │ │ Co-financed │ │ For profit │
└──────────┘ └────────────┘ └──────────────┘ └────────────┘ └────────────┘
```

Figure 3: *Funding sources.*

To further assess methodological quality, we will note ("Yes" or "No") whether a flow-chart is publicly available.

Single or multicentre status will be determined through review [9]. A trial will be considered a multicentre trial if more than one centre is involved. In case of missing information, trials will be classified as multicentre when there is reporting of both several ethics committees and different affiliations of authors. On the other hand, if the report stated only a single ethics committee and a single author affiliation, the trial will be classified as a single centre, unless other information calls for multicentre.

The geographical trial setting will be noted based on the continents participating in the trial: North America, Europe, South America, Asia, Africa and Australia. All the continents involved in the study will be registered.

RoB item	Low RoB	High RoB
Funding	No funding and not-for-profit funding will be considered as low RoB.	For-profit funding and co-financed funding will be considered high RoB. If funding is not reported, it will also be considered high RoB.
Conflict of interest	If conflict of interest is reported as "none," it will be considered low RoB.	If conflict of interest is not reported as "none," it will be considered high RoB.
Flowchart	If a flowchart is publicly available, it will be considered low RoB.	If a flowchart is not publicly available, it will be considered high RoB.
Number of study locations	It will be considered low RoB if more than one centre participated in the trial.	It will be considered high RoB if only one centre participated in the trial, or if it is unclear how many centres participated.

Table 4: *Risk of bias (RoB) components currently not included in the Cochrane RoB tool.*

Two reviewers will independently evaluate eligibility, risk of bias, assessment of ITT/mITT, and handling of missing data. Disagreements will be resolved by discussion until consensus.

Data synthesis

Data synthesis will assess ACR20 data at primary endpoint. We will present differences among different strata by estimating the ratio of odds ratios (ROR). We will assume that the observed number of responders in each arm of each trial has a binomial distribution. Accordingly, intervention effects will be modelled as log-odds ratios and outcomes will be coded so that ORs > 1 correspond to beneficial intervention effects. We will estimate the odds ratio (OR) of trials with the given characteristic using random-effects meta-analyses. For each drug corresponding to a meta-analysis, we will derive the difference between pooled estimates from trials with different trial characteristics (e.g., different ITT approaches). Formal tests of interaction between ITT analysis and estimated treatment benefits (active compared to control) will be performed using the following statistical methodology. After identifying a all trials available for the different biological and targeted agents, we will record characteristics of individual studies (ITT: y/n; Type of ITT: ITT/mITT/AO/PP/Other/Unclear; Type of modification: Treatment/Post baseline assessment/Combined/Other; Handling of missing data in the trial: LOCF/BOCF/Other/Unclear;

Handling of NRI: All wd/Lack Of Efficacy wd/Other wd/Unclear) and compare treatment effects within each biological or targeted agent.

We will fit *empirical Bayesian* hierarchical bias models using the generalized linear mixed models (GLMM) [56]. Mean intervention effects may differ among trials with and without the reported study characteristic. Variation in bias among trials within biologics or targeted agents' trials is quantified and adjusted for with a fixed factor in the model. The GLMMs, like linear mixed models, assume normal (Gaussian) random effects. Conditional on these random effects, data can have any distribution in the exponential family. The exponential family comprises many of the elementary discrete and continuous distributions. The analyses will be performed using the GLIMMIX procedure in SAS (v. 9.2; SAS Institute Inc., Cary, NC, USA) [57]. The syntax is similar to that of the MIXED procedure and includes CLASS, MODEL, and RANDOM statements. Using the GLIMMIX procedure, we will perform mixed-effects logistic regression with an arm-based, random-effects model within an *empirical Bayes* framework:

```
Proc Glimmix;
Class Trial Drug Group Characteristic;
Model Counts/Total = Group Characteristic Group*Characteristic / Solution;
Random Trial Trial*Group;
Lsmeans Group*Characteristic / cl ilink;
Run;
```

The PROC GLIMMIX statement invokes the procedure. The CLASS statement instructs the procedure to treat the variables Trial, Drug, Group, and Characteristic as classification variables. The MODEL statement specifies the response variable as a sample proportion using the r/N syntax: Counts/Total corresponds to Y_{iA}/N_{iA} for observations from Group A and to Y_{iB}/N_{iB} or observations from Group B. The SOLUTION option in the MODEL statement requests a listing of the fixed-effects parameter estimates. Because of the response/N syntax, the GLIMMIX procedure defaults to the binomial distribution, with the default logit link. The RANDOM statement specifies that the linear predictor contains intercept terms that randomly vary at the level of the Trial and Trial×Group effects. The default estimation technique in GLMMs is residual pseudo-likelihood (RSPL) with a

subject-specific expansion. The default optimization technique for GLMMs is the Quasi-Newton method. Because a residual likelihood technique is used to compute the objective function, only the covariance parameters are participating in the optimization.

The LsMeans statement requests the least-squares means of the interaction between group (active vs. control) and the individual study characteristic effect on the logit scale. The CL option requests their confidence limits. The ILINK option adds estimates, standard errors, and confidence limits on the mean (probability) scale.

DISCUSSION

Biased results from RCTs ultimately put the patients at risk for being treated with pharmaceuticals with questionable efficacy and which may cause harm. Taking into account the expenses of accompanying RA treatment, this study is not only biomedical but also a socioeconomic necessity.

The term mITT is used to describe different methods for excluding participants post-randomised from analysis, thereby affecting and disregarding not only the ITT principle but also—and more importantly—the overriding purpose of ITT. Post-randomisation exclusions are known to induce bias, and theoretically mITT will introduce bias [21;22]. Our study aim to establish if the bias is of practical concern, and focuses on the direction and magnitude of bias associated with mITT analyses. This study will present arguments as to why mITT approximates ITT or point to the problems concerning the use of mITT. As the term mITT embraces a broad notion of trials, we will delve into how the different types of modification influence effect size. This study may come out with neutral findings—which would not imply that overall bias associated with mITT analyses can be excluded, but may indicate that our study lacks the statistical power necessary to detect the bias. If some form of mITT can substitute ITT, guidelines regarding the use of mITT should be issued.

Our primary objective is to examine whether mITT is associated with different effect sizes, implying empirical evidence for bias in treatment effects. ITT prevents attrition bias when evaluating treatment assignment but may not provide a true estimate of treatment effect if some patients are non-adherent [15]. As the term "bias" comprises deviation from the true intervention effect, it can be perceived as misleading to regard systematic errors in treatment effect between mITT and ITT analyses as "bias," given that ITT analysis may fail to provide a true evaluation of the intervention effect. However, ITT analysis is recommended as the least biased way to estimate intervention effects [7] and concerns regarding the systematic errors between mITT and ITT remain, regardless of terminology.

This study may point to potential bias and disadvantages in the handling of missing data in RCTs, otherwise known for having a low risk of bias compared with other study designs [58]. SI has been criticised on a theoretical level, but its implication on efficacy outcomes in RA trials is uncharted. Accordingly, this study may provide empirical evidence that can support or contradict existing critics. The study examines potential bias associated with industry funding. It may prove difficult to assess bias, as most RA trials concerning biological and targeted agents have some degree of industry input. Being unable to reject industry bias and unable to estimate the influence, direction, and magnitude of such, the validity of trial results in this industry-permeated field of research is open to conjecture [14].

Dissemination

First author Anna Dossing will draft a paper describing the systematic review; the meta-epidemiological study will be disseminated by peer-review publication and conference presentations.

HISTORY

Protocol first published: 11. December 2013

CONTRIBUTION OF AUTHORS

All authors participated in the conception and design of this protocol. RC provided statistical advice for the design and analysis. AD drafted the manuscript. All authors critically reviewed the manuscript and approved the final version.

FUNDING

This research received grants from the Michaelsen foundation. No sponsor was involved in study design, and no sponsor will have authority in the collection, management, analysis, and interpretation of data. Writing of the report and the decision to submit the results for publication is strictly made by the authors. Musculoskeletal Statistics Unit, The Parker Institute, is supported by grants from the Oak Foundation. The Copenhagen Trial Unit is funded by the Danish state.

DECLARATION OF INTEREST

The authors declare no conflicts of interest.

ABBREVIATIONS

ACR	American College of Rheumatology
AO	As observed
bDMARDs	Biological disease-modifying anti-rheumatic drugs
BOCF	Baseline Observation Carried Forward
CRP	C-reactive protein
csDMARDs	Conventional synthetic disease-modifying anti-rheumatic drugs
DAS28	European Disease Activity Score
DMARDs	Disease-modifying anti-rheumatic drugs
EMA	European Medicines Agency
ESR	Erythrocyt sedimentation rate
EULAR	European League Against Rheumatism

FDA	US Food and Drug Administration
GCP	Good clinical practice
GLMM	Generalized linear mixed models
ICTRP	WHO Clinical trials Portal
IR	Inadequate response
ITT	Intention-To-Treat
LOCF	Last Observation Carried Forward
MAR	Missing at random
MCAR	Missing completely at random
MI	Multiple imputation
mITT	modified Intention-To-Treat
MNAR	Missing not at random
MTX	Methotrexat
NRI	Non-responder imputation
OR	Odds ratio
PP	Per protocol
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RA	Rheumatoid Arthritis
RCT	Randomised controlled trial
RF	Rheumatoid factor
RoB	Risk of bias
ROR	Ratio of odds ratios
RSPL	Residual pseudo-likelihood
sDMARDs	Synthetic disease-modifying anti-rheumatic drugs
SI	Single imputation
tsDMARD	Targeted synthetic disease-modifying anti-rheumatic drugs
wd	Withdrawal patients

REFERENCES

(1) Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ 2001; 323(7303):42-46.

(2) Sterne JA, Juni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. Stat Med 2002; 21(11):1513-1524.

(3) Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008; 336(7644):601-605.

(4) Savovic J, Jones H, Altman D, Harris R, Juni P, Pildal J et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. Health Technol Assess 2012; 16(35):1-82.

(5) Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011; 343:d5928.

(6) Gluud LL. Bias in clinical intervention research. Am J Epidemiol 2006; 163(6):493-501.

(7) Higgins J, Green S, (editors). Cochrane handbook for Systematic Reviews of Interventions. Version 5.1.0. 2011. The Cochrane Collaboration. Ref Type: Serial (Book,Monograph)

(8) Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med 2001; 135(11):982-989.

(9) Dechartres A, Boutron I, Trinquart L, Charles P, Ravaud P. Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study. Ann Intern Med 2011; 155(1):39-51.

(10) Bafeta A, Dechartres A, Trinquart L, Yavchitz A, Boutron I, Ravaud P. Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. BMJ 2012; 344:e813.

(11) Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ 2013; 346:f2304.

(12) Montedori A, Bonacini MI, Casazza G, Luchetta ML, Duca P, Cozzolino F et al. Modified versus standard intention-to-treat reporting: are there differences in methodological quality, sponsorship, and findings in randomized trials? A cross-sectional study. Trials 2011; 12:58.

(13) Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 2003; 290(7):921-928.

(14) Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev 2012; 12:MR000033.

(15) Shrier I, Steele RJ, Verhagen E, Herbert R, Riddell CA, Kaufman JS. Beyond intention to treat: What is the right question? Clin Trials 2013.

(16) Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say? Clin Trials 2007; 4(4):350-356.

(17) Kruse RL, Alper BS, Reust C, Stevermer JJ, Shannon S, Williams RH. Intention-to-treat analysis: who is in? Who is out? J Fam Pract 2002; 51(11):969-971.

(18) Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999; 319(7211):670-674.

(19) Baron G, Boutron I, Giraudeau B, Ravaud P. Violation of the intent-to-treat principle and rate of missing data in superiority trials assessing structural outcomes in rheumatic diseases. Arthritis Rheum 2005; 52(6):1858-1865.

(20) Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: systematic review. BMJ 2010; 340:c2697.

(21) Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. Int J Epidemiol 2005; 34(1):79-87.

(22) Nuesch E, Trelle S, Reichenbach S, Rutjes AW, Burgi E, Scherer M et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. BMJ 2009; 339:b3244.

(23) Altman DG, Bland JM. Missing data. BMJ 2007; 334(7590):424.

(24) Boers M. Missing data in trials: do we have to keep carrying the last observation forward? Arthritis Rheum 2008; 59(1):2-3.

(25) EMA. Guideline on missing data in confirmatory clinical trials. London: European Medicines Agency (EMA); 2010.

(26) White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ 2011; 342:d40.

(27) Shao J, Jordan DC, Pritchett YL. Baseline observation carry forward: reasoning, properties, and practical issues. J Biopharm Stat 2009; 19(4):672-684.

(28) Molnar FJ, Hutton B, Fergusson D. Does analysis using "last observation carried forward" introduce bias in dementia research? CMAJ 2008; 179(8):751-753.

(29) Altman DG. Missing outcomes in randomized trials: addressing the dilemma. Open Med 2009; 3(2):e51-e53.

(30) Baron G, Boutron I, Giraudeau B, Ravaud P. Reporting of radiographic methods in randomised controlled trials assessing structural outcomes in rheumatoid arthritis. Ann Rheum Dis 2007; 66(5):651-657.

(31) Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009; 338:b2393.

(32) FDA. Guidance for Industry, Clinical Development Programs for Drugs, Devices, and Biological Products for the Treatment of Rheumatoid Arthritis (RA). 1999.

(33) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31(3):315-324.

(34) Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995; 38(1):44-48.

(35) Nielen MM, van SD, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004; 50(2):380-386.

(36) Dahlqvist SR. Rheumatoid arthritis increased the risk for myocardial infarction in women. ACP J Club 2003; 139(2):50.

(37) Scott DL, Symmons DP, Coulton BL, Popert AJ. Long-term outcome of treating rheumatoid arthritis: results after 20 years. Lancet 1987; 1(8542):1108-1111.

(38) Welsing PM, van Gestel AM, Swinkels HL, Kiemeney LA, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum 2001; 44(9):2009-2017.

(39) van DH, van AJ, Lard LR, Visser K, Ronday HK, Hulsmans HM et al. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. Arthritis Rheum 2007; 56(5):1424-1432.

(40) Smolen JS, Landewe R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. Ann Rheum Dis 2010; 69(6):964-975.

(41) Furst DE, Keystone EC, Braun J, Breedveld FC, Burmester GR, De BF et al. Updated consensus statement on biological agents for the treatment of rheumatic diseases, 2011. Ann Rheum Dis 2012; 71 Suppl 2:i2-45.

(42) Tugwell P, Singh JA, Wells GA. Biologicals for rheumatoid arthritis. BMJ 2011; 343:d4027.

(43) Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis 2013.

(44) Strand V, Sokolove J. Randomized controlled trial design in rheumatoid arthritis: the past decade. Arthritis Res Ther 2009; 11(1):205.

(45) Stein CM, Pincus T. Placebo-controlled studies in rheumatoid arthritis: ethical issues. Lancet 1999; 353(9150):400-403.

(46) EMA. Guideline on clinical investigation of medicinal products other than NSAIDs for treatment of rheumatoid arthritis. 2011.

(47) FDA. Guidance for Indusry, Rheumatoid Arthritis: Developing Drug Products for Treatment. 2013.

(48) Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995; 38(6):727-735.

(49) EMA. Points to consider on clinical investigation of medicinal products other than NSAIDS for treatment of rheumatoid arthritis. London: European Medicines Agency (EMA); 2003.

(50) EMA. Draft guideline on clinical investigation of medicinal products other than non-steroidal anti-inflammatory drugs for treatment of rheumatoid arthritis. London: European Medicines Agency (EMA); 2011.

(51) Aletaha D, Landewe R, Karonitsch T, Bathon J, Boers M, Bombardier C et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Arthritis Rheum 2008; 59(10):1371-1377.

(52) Dossing A, Tarp S, Furst DE, Gluud C, Wells GA, Beyene J et al. Attrition bias in rheumatoid arthritis randomised trials with different modified intention-to-treat approaches: a meta-epidemiological study. PROSPERO 2013; CRD42013006702.

(53) Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med 2009; 151(4):W65-W94.

(54) Wells G, Becker JC, Teng J, Dougados M, Schiff M, Smolen J et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response

criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009; 68(6):954-960.

(55) Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med 2012; 157(6):429-438.

(56) Platt RW, Leroux BG, Breslow N. Generalized linear mixed models for meta-analysis. Stat Med 1999; 18(6):643-654.

(57) Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. Cochrane Database Syst Rev 2009;(4):CD007848.

(58) Jakobsen J C & Gluud C. The Necessity of Randomized Clinical Trials. British Journal og Medicine and Clinical Research 2013; 3(4):1453-1468.

APPENDIX 1

The Cochrane Library search strategy:

ID	Search
#1	MeSH descriptor: [Recombinant Fusion Proteins] explode all trees
#2	MeSH descriptor: [Antibodies, Monoclonal] explode all trees
#3	MeSH descriptor: [Receptors, Tumor Necrosis Factor] explode all trees
#4	MeSH descriptor: [Receptors, Interleukin-1] explode all trees
#5	MeSH descriptor: [Receptors, Interleukin-6] explode all trees
#6	MeSH descriptor: [Monokines] explode all trees
#7	monoclonal antibody ca2
#8	TNFR-Fc fusion protein
#9	MeSH descriptor: [Interleukin 1 Receptor Antagonist Protein] explode all trees
#10	etanercept
#11	enbrel
#12	infliximab
#13	remicade
#14	adalimumab
#15	humira
#16	D2E7
#17	anakinra
#18	kineret
#19	antril
#20	abatacept
#21	CTLA4Ig
#22	orencia
#23	rituximab
#24	rituxan
#25	idec c2b8
#26	golimumab
#27	simponi
#28	cnto-148
#29	tocilizumab
#30	atlizumab
#31	actemra
#32	roactemra
#33	certolizumab
#34	CDP870
#35	cimzia
#36	"TNFR:Fc":ti,ab,kw (Word variations have been searched)
#37	tofacitinib:ti,ab,kw (Word variations have been searched)

#38	MeSH descriptor: [Janus Kinases] explode all trees
#39	Xeljanz:ti,ab,kw (Word variations have been searched)
#40	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39
#41	MeSH descriptor: [Arthritis, Rheumatoid] explode all trees
#42	Rheumatoid:ti or Rheumatoid:ab (Word variations have been searched)
#43	arthriti*:ti or arthriti*:ab (Word variations have been searched)
#44	#42 and #43
#45	#41 or #44
#46	#40 and #45 in Trials

EMBASE search strategy:

1	abatacept.mp.
2	adalimumab.mp.
3	certolizumab.mp.
4	etanercept.mp.
5	CDP870.mp.
6	golimumab.mp.
7	infliximab.mp.
8	rituximab.mp.
9	tocilizumab.mp.
10	humira.mp.
11	trudexa.mp.
12	orencia.mp.
13	cimzia.mp.
14	enbrel.mp.
15	simponi.mp.
16	rituxan.mp.
17	mabthera.mp.
18	actemra.mp.
19	RoActemra.mp.
20	monoclonal antibodies.mp. or exp Antibodies, Monoclonal/
21	exp Monokines/

22	exp Receptors, Interleukin-1/
23	exp Receptors, Interleukin-6/
24	exp Polyethylene Glycols/
25	exp Immunoglobulin G/
26	exp Immunoconjugates/
27	immunoglobulin fab fragments.mp. or exp Immunoglobulin Fab Fragments/
28	t-lymphocytes.mp. or exp T-Lymphocytes/
29	exp tumor necrosis factor inhibitor/
30	exp interleukin 1 receptor blocking agent/
31	D2E7.mp.
32	anakinra.mp.
33	kineret.mp.
34	antril.mp.
35	CTLA4Ig.mp.
36	idec c2b8.mp.
37	cnto-148.mp.
38	atlizumab.mp.
39	tofacitinib.mp.
40	exp Janus kinase inhibitor/
41	*tumor necrosis factor receptor/dt [Drug Therapy]
42	or/1-41
43	exp Random Allocation/
44	exp Single-Blind Method/
45	exp Double-Blind Method/
46	Placebo.mp.
47	Randomi?ed controlled trial$.mp.
48	rct.mp.
49	Random allocation.mp.
50	Randomly allocated.mp.
51	Allocated randomly.mp.
52	(allocated adj2 random).mp.
53	Single blind$.mp.
54	Double blind$.mp.
55	((treble or triple) adj blind$).mp.
56	Placebo$.mp.

57	or/43-56
58	rheumatoid.ti,ab.
59	*rheumatoid arthritis/
60	58 or 59
61	42 and 57 and 60
62	limit 61 to (book or book series or conference abstract or conference paper or conference proceeding or "conference review")
63	61 not 62

LILACS search strategy:

(tw:(tw:((tw:(rheumatoid)) AND (tw:(etanercept)) OR (tw:(enbrel)) OR (tw:(infliximab)) OR (tw:(remicade)) OR (tw:(adalimumab)) OR (tw:(humira)) OR (tw:(d2e7)) OR (tw:(anakinra)) OR (tw:(kineret)) OR (tw:(antril)) OR (tw:(abatacept)) OR (tw:(ctla4ig)) OR (tw:(orencia)) OR (tw:(rituximab)) OR (tw:(rituxan)) OR (tw:(idec c2b8)) OR (tw:(golimumab)) OR (tw:(simponi)) OR (tw:(cnto-148)) OR (tw:(tocilizumab)) OR (tw:(atlizumab)) OR (tw:(actemra)) OR (tw:(roactemra)) OR (tw:(certolizumab)) OR (tw:(cdp870)) OR (tw:(cimzia)) OR (tw:(tnfr:fc)) OR (tw:(tofacitinib)) OR (tw:(janus kinases)) OR (tw:(xeljanz))) AND db:("LILACS"))) AND (tw:(random*))

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	8
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	8
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	9
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	9+supplement
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	10+13
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	11+13
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	14-17
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	17
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	17-19

PRISMA 2009 Checklist

Page 1 of 2

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	14-17
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	17-19
RESULTS			Study protocol, no results
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	19
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Study protocol, no results
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	21

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097.*

PRISMA 2009 Checklist

doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Page 2 of 2

BMJ Open

Interpreting trial results following use of different intention-to-treat approaches for preventing attrition bias: A meta-epidemiological study protocol

SCHOLARONE™
Manuscripts

Interpreting trial results following use of different intention-to-treat approaches for preventing attrition bias: A meta-epidemiological study protocol

Anna Dossing[1,2], Simon Tarp[1], Daniel E. Furst[3], Christian Gluud[4], Joseph Beyene[5], Bjarke B. Hansen[1], Henning Bliddal[1], Robin Christensen[1]

Affiliations
1: Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospitals, Bispebjerg and Frederiksberg, Denmark
2: Department of Clinical Medicine, Faculty of Medical and Health Sciences, University of Copenhagen, Denmark
3: University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, USA
4: Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
5: Department of Clinical Epidemiology & Biostatistics, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

Correspondence to:
Dr. Robin Christensen, MSc, PhD
Head of Musculoskeletal Statistics Unit;
The Parker Institute, Department of Rheumatology.
Copenhagen University Hospital, Bispebjerg and Frederiksberg.
Nordre Fasanvej 57
DK-2000 Copenhagen Frederiksberg
Denmark
Phone: +45 3816 4165
Fax: +45 3816 4159
E-Mail: Robin.Christensen@frh.regionh.dk

Keywords: Rheumatoid Arthritis, intention to treat analysis, meta-analysis, bias, epidemiologic.

Word count: (excluding title page, abstract, references, figures and tables): 4,654

ABSTRACT

Introduction: When subjects drop out of randomised clinical trials, as frequently happens, the intention-to-treat principle does not apply, potentially leading to attrition bias. Data lost from patient dropout/lack of follow-up are statistically addressed by imputing, a procedure prone to bias. Deviations from the original definition of ITT are referred to as modified ITT (mITT). As yet, the impact of the potential bias associated with mITT has not been assessed. Our objective is to investigate potential bias and disadvantages of performing mITT and evaluate possible concerns when executing different mITT approaches in meta-analyses.

Methods and analysis: Using meta-epidemiology on randomised trials considered less prone to bias (i.e., good internal validity) and assessing biological or targeted agents in patients with rheumatoid arthritis (RA), we will meta-analyse data from 10 biological and targeted drugs based on collections of trials that would correspond to 10 individual meta-analyses.

Ethics and dissemination: This study will enhance transparency for evaluating mITT treatment effects described in meta-analyses. The intended audience will include health care researchers, policymakers and clinicians. Results of the study will be disseminated by peer-review publication.

Protocol registration: In PROSPERO CRD42013006702, 11. December 2013

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This meta-epidemiological study is the first to focus on bias associated with mITT and its impact on effect size.

- This study will influence how results from RCTs and meta-analysis should be interpreted; this study will provide a framework for designing, conducting and reporting future RCTs in order to minimise attrition bias.

- In practice, various definitions of both ITT and mITT are used. Some may argue that our definitions are not stringent enough or inverse overly stringent.

INTRODUCTION

Meta-epidemiological research

Inadequate quality of trials may distort the results from meta-analyses and systematic reviews (1). Consequently, meta-epidemiological studies are carried out to quantitatively evaluate bias across many randomised clinical trials (RCTs) in different meta-analyses (2-4). Overall flaws in the design, conduct, analysis, and reporting of RCTs can introduce bias—systematic errors that lead to overestimating or underestimating the benefits and the harms of treatment (5;6).

Methodological quality and risk of bias assessment

The Cochrane Collaboration's Handbook for Systematic Reviews of Interventions provides guidance to authors to critically review trial outcome using the risk of bias (RoB) assessment tool (5). The RoB tool requires authors to evaluate the well-established strengths and limitations of RCTs, including sequence generation, allocation concealment, blinding of participants, personnel and outcome assessors, loss to follow-up with failure to apply the intention-to-treat (ITT) principle, and selective outcome reporting (5;7). In the course of meta-epidemiological studies, other sources of bias in RCTs have been identified, such as significant discrepancies favouring intervention in single (rather than multicentre) trials, in trials with small (rather than large) sample sizes, and in using subjective (rather than objective) outcome measures (3;8-11). Most recently, funding source has become a distinct possibility as a source of bias, with for-profit organisation funding likely favouring pro-intervention results (12-14), however there is an on-going debate as to whether funding should be regarded as a RoB item (15;16).

The intention-to-treat principle

The ITT principle has two main rules: 1) it requires all participants in a RCT to be analysed according to their original allocation, regardless of their adherence to the trial protocol; and 2) all randomised participants must be included in the analysis. The first rule acts to conserve randomisation, which is executed to avoid selection bias and thereby produces treatment groups in which the distributions of prognostic factors—both known and unknown—are similar. The second rule serves to avoid attrition bias when evaluating a treatment assignment(17). Attrition bias is attributed to systematic differences between groups in withdrawals from the study (7).

Frequently, not all participants in RCTs are analysed as they were initially randomised (18-21). Various deviations from the definition of ITT have led to the vaguely defined term "modified ITT" (mITT), which may compromise true randomisation. The incidence of trials reporting the use of mITT has increased over the years (22). Trials that report mITT exclude patients from analysis post-randomisation compared to trials reporting on the ITT population (12). Although post-randomisation exclusion is known to produce bias (23;24), the potential magnitude and direction of bias associated with mITT is unknown.

Based on the most common deviations from ITT described in the literature (19;20;22), mITT can be divided into four categories, as illustrated in **Figure 1**.

Handling of missing data

Executing ITT analysis can be difficult, as missing data is common (25;26). Missing data also affects mITT analysis unless mITT is defined by a complete case analysis. Missing data comprises single missing data points as well as missing datasets due to withdrawal (26). Data can be missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). MNAR occurs when missing data depend on unobserved values: and the missing data may lead to bias (25-27). Regardless of trial design and objective, it is likely that some of the data will be MNAR.

Problems from missing data in RCTs can be somewhat mitigated by data imputation—a procedure by which missing data are replaced with a statistically founded estimate (27;28). Two single imputation (SI) methods often used are "*last observation carried forward"* (LOCF) or "*baseline observation carried forward"* (BOCF). LOCF uses the last observed value in place of the missing outcome, whereas BOCF uses the baseline observation as the value replacing the outcome (27;29). LOCF and BOCF have been widely criticised, (21;30;31) and the use of SI risk introducing bias, e.g. by minimising variation (27;32). A newer and more promising imputation method is multiple imputation (MI), which involves creating several different plausible imputed datasets based on a

Bayesian approach and then combining the results obtained from each of them, hence reducing the risk of underestimating variance (27;32;33).

Loss to follow-up is frequent in RCTs and can be attributed to a variety of causes (23;25). Most patients drop out of trials due to lack of efficacy, adverse effects of treatment, or both (26). When handling a dichotomous outcome in a responder analysis, it is often considered suitable to attribute patients' withdrawal to lack of efficacy, and therefore assume treatment failure (27;34) (also referred to as non-responder imputation (NRI)). NRI is applied inconsistently to different withdrawal populations, e.g. all patients who withdraw are considered treatment failures as opposed to only patients who withdraw due to lack of efficacy are considered treatment failures. In principle there should be no objection to applying NRI to different withdrawal populations (27) and no empirical evidence have documented whether it affects trial results.

Rationale for this meta-epidemiological study

Our extensive search did not find any previous systematic assessment of empirical evidence for bias associated with mITT (23;24). Furthermore, mITT has not been assessed as to whether the type or number of modifications applied affects the estimated efficacy outcome differently. To investigate potential bias, we focus on trials assessing biological (or targeted) interventions in patients with rheumatoid arthritis (RA), as these trials are relatively recent and we therefore anticipate that these yield reasonable internal validity. With regards to imputation, we wish to examine how imputation of missing data affects RA trials investigating clinical response, as this have, to our knowledge, not previously been done. The study also aims to shed a light on the effect of selectively applying NRI to different withdrawal populations, which so far has remained unexplored.

Biological and targeted agents for rheumatoid arthritis

RA is a chronic inflammatory autoimmune disease characterised by joint swelling and joint tenderness with destruction of synovial joints primarily affecting the hands (35-38). The inflammatory load drives the destructive progression of the disease and leads to severe disability and premature mortality (39;40). Diagnosis of RA requires that patients have a minimum of four

criteria that persist for 6 months (35). The seven qualifying criteria are: 1) morning stiffness; 2) arthritis of three or more joints or joint areas; 3) arthritis of hand joints; 4) symmetric arthritis; 5) rheumatoid nodules; 6) serum rheumatoid factor (RF); and 7) radiographic changes.

Treatment of RA encompasses multiple interventions. Though RA has been shown to progress over time despite treatment (39), early therapeutic intervention improves clinical outcomes and reduces the accrual of joint damage and disability (41). RA can be managed using disease-modifying anti-rheumatic drugs (DMARDs). DMARDs form two major classes: synthetic chemical compounds (sDMARDs) and biological agents (bDMARDs). sDMARDs comprise the conventional DMARDs (csDMARDs) (e.g., methotrexat [MTX] and sulfasalazin) and the new targeted sDMARDs (tsDMARDs); e.g., tofacitinib). bDMARDs is a heterogeneous group of pharmaceuticals including abatacept, adalimumab, anakinra certolizumab, etanercept, golimumab, infliximab, rituximab, and tocilizumab that help to control the autoimmune inflammation associated with arthritis (42-45). The DMARDs of interest are summarised in **Table 1** (45).

DMARDs	Pharmaceutical
csDMARDs	methotrexat, sulfosalazin, leflunomid, hydroxychloroquine
tsDMARD	Tofacitinib
bDMARDs	adalimumab, abatacept, anakinra, certolizumab pegol, etanercept, golimumab, infliximab, rituximab, tocilizumab

Table 1*: Overview of disease-modifying anti-rheumatic drugs (DMARD) groups: conventional DMARDs (csDMARDs); the new targeted sDMARDs (tsDMARDs); and biological agents (bDMARDs).*

In RA RCTs on biological and targeted agents, the control group is commonly treated with MTX, as lack of treatment can lead to irreversible loss of physical function in RA patients (46;47). Rescue therapy (e.g., regulation of dose or addition of MTX or bDMARD or tsDMARD) is acknowledged by both the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA) and is typically offered when treatment response is evaluated as inadequate (48;49). Because patients who receive rescue therapy are commonly encountered as withdrawal patients in the trials protocols; rescue therapy challenges the ITT principle and may contribute to attrition bias.

The treatment effect in RA RCTs is typically evaluated using the American College of Rheumatology (ACR) response criteria and/or variations of the European Disease Activity Score (DAS28) (46). The DAS28 score discriminates between high and low disease activity (36). The ACR response criteria's definition of improvement in RA trials includes improvement in the joint counts and improvement in at least three of the five following: 1) patient assessment; 2) physician assessment; 3) erythrocyte sedimentation rate; 4) pain scale; and 5) functional questionnaire (50): Improvement is described as ACR20, ACR50 or ACR70 and refers to a 20%, 50%, or 70% improvement in tender and swollen joint count and 20%, 50%, or 70% improvement in three of the five core set measures listed above. The EMA, the European League Against Rheumatism (EULAR), and the ACR consent that validated composite clinical outcomes such as the ACR response criteria should be used to document efficacy of treatment. Specifically the ACR20, ACR50, or ACR70 should be used to document signs and symptoms after 3–6 months (51-53).

Objective

Our primary objective is to examine

1. Whether mITT is associated with different effect sizes compared to ITT.

Secondarily we wish to examine

2a. How the choice of imputation technique influences the effect size.

2b. How different ITT populations affect effect size.

Hypotheses

We hypothesise that different ITT populations are associated with differences in treatment effect sizes. We expect the level of statistical significance will depend on the type of modification, the number of modifications applied, and the percentage of patients excluded from final efficacy analysis. Furthermore, we hypothesise that the use of different mITT approaches is strongly associated with industry funding (12) and these modifications are applied to artificially increase the effect size for the apparent benefit in trials directly related to a specific pharmaceutical company.

METHODS AND ANALYSIS

Protocol and registration

Our protocol is registered on PROSPERO (CRD42013006702)(54); our manuscript conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for reporting systematic reviews and meta-analyses (55).

Eligibility criteria

All RA RCTs concerning EMA- or FDA-approved biologics and targeted agents evaluating efficacy as ACR20, ACR50, or ACR70 will be considered eligible, independent of whether they include ITT or mITT analysis. We will include published RCTs. Open-label studies will be excluded from our analysis because performance and detection bias are inherent when the outcome is not considered an objective endpoint (3).

Search and selection of trials and meta-analyses

We will search PubMed, EMBASE, The Cochrane Central Register of Controlled Trials (CENTRAL), and LILACS using a combination of keywords and text words related to RA. The PubMed search strategy is listed below. See **Appendix 1** for the EMBASE, the Cochrane Library, and LILACS search strategies. The World Health Organisation (WHO) Clinical trials Portal (ICTRP), clinicaltrials.gov, FDA, EMA and pharmaceutical companies' trial result databases will be searched to identify unpublished data.

PubMed search strategy: (((("Receptors, Tumor Necrosis Factor"[nm] OR TNFR:Fc OR "TNFR-Fc fusion protein"[Supplementary Concept] OR "TNFR-Fc fusion protein"[All Fields] OR "etanercept"[All Fields] OR "enbrel"[All Fields]) OR ("infliximab"[Supplementary Concept] OR "infliximab"[All Fields] OR "remicade"[All Fields] OR "mab ca2"[All Fields] OR "monoclonal antibody ca2"[All Fields]) OR ("adalimumab"[Supplementary Concept] OR "adalimumab"[All Fields] OR "humira"[All Fields]) OR ("interleukin 1 receptor antagonist protein"[MeSH Terms] OR "interleukin 1 receptor antagonist protein"[All Fields] OR "anakinra"[All Fields] OR "kineret"[All Fields] OR "antril"[All Fields]) OR ("abatacept"[Supplementary Concept] OR "abatacept"[All Fields]) OR CTLA4Ig[All Fields] OR "orencia"[All Fields]) OR ("rituximab"[Supplementary Concept] OR

"rituximab"[All Fields] OR "rituxan"[All Fields] OR "idec c2b8"[All Fields]) OR ("golimumab"[All Fields] OR "golimumab"[Supplementary Concept] OR "simponi"[All Fields] OR "cnto-148"[All Fields] OR ("cnto"[All Fields] AND "148"[All Fields])) OR ("tocilizumab"[All Fields] OR "tocilizumab"[Supplementary Concept] OR "atlizumab"[All Fields] OR "actemra"[All Fields]) OR ("certolizumab"[All Fields] OR "certolizumab pegol"[Supplementary Concept] OR "CDP870"[All Fields] OR ("cdp"[All Fields] AND "870"[All Fields]) OR "cimzia"[All Fields]) OR ("tofacitinib"[Supplementary Concept] OR "tofacitinib"[All Fields]) OR ("Antibodies, Monoclonal"[Mesh] OR "Monokines"[Mesh] OR "Receptors, Interleukin-1"[Mesh] OR "Receptors, Interleukin-6"[Mesh])) AND ("Randomized Controlled Trial"[ptyp] OR "Controlled Clinical Trial"[ptyp] OR "Multicenter Study"[ptyp] OR "randomized"[tiab] OR "randomised"[tiab] OR "placebo"[tiab] OR "randomly"[tiab] OR "trial"[tiab] OR randomized controlled trials[mh] OR random allocation[mh] OR double-blind method[mh] OR single-blind method[mh]) AND ("Arthritis, Rheumatoid"[MeSH Terms] OR (Rheumatoid[text word] AND arthriti*[text word]))) NOT (animals[mh] NOT human[mh]).

Date for final database update will be stated.

Data extraction

All RCT publications will be assigned an ID, and we will extract information on author, year of publication, journal of publication, company study name, company study number, and registration number. All characteristics will be typed into a custom made database, RHEUMATRIALS. We will extract baseline characteristics and inclusion criteria. RA will be classified on the basis of mean disease duration as early (≤ 6 months), established (6 months to 2 years), or late (> 2 years). Study duration until registration of primary outcome and duration of the longest placebo-controlled period will be extracted. We will extract type of primary outcomes (i.e., safety, efficacy, or both) and the name of the primary outcome (e.g., ACR20).

Trials will be classified based on the included patient populations. We will distinguish among three possibilities: whether patients have exhausted all csDMARD treatment options, whether they have

had inadequate response to csDMARD treatment and are candidates for biologics or a targeted agent, or whether they have experienced an inadequate response to a bDMARD or tsDMARD.

Regarding the intervention, we will extract the main drug of interest, the allocation, the form of administration, dosing of the drug, and registration of whether dose is equivalent to standard dose. We will register if treatment with csDMARD concomitant to the intervention was allowed. Offering of rescue will be registered as "yes" or "no": time-point and the content of rescue will also be registered.

We will extract the total number of patients reaching the primary outcome and the total number of patients in the analysis. For efficacy outcomes, we will assess ACR20, ACR50, and ACR70. For sensitivity , we will extract the continuous efficacy outcome DAS28 as the mean change from baseline in each intervention group with the corresponding dispersion, when available. If DAS28 was reported based on C-reactive protein (CRP) and on erythrocyt sedimentation rate (ESR), we will extract only DAS28 (ESR), as both outcomes are in aggreement (56). For safety outcomes, we will assess overall withdrawals, withdrawals due to adverse events, and number of serious adverse events.

Intention-to-treat analysis on efficacy outcome

We will extract the number of patients randomised in the trial in total, the number of randomised patients per intervention group, the number of patients completing each intervention group, and the number of patients included in efficacy analysis per intervention group. The method of efficacy analysis will be extracted on two levels. The first level will assess the *reported* protocolised method of analysis. The second level will register the *applied* method of analysis. The analysis population will be categorised as: "ITT," "mITT," "as observed," "per protocol," "other," or "unclear." Categorisation will be based on the population included in ACR20 efficacy analysis. A full ITT analysis will comprise both patients with clinically assessed outcome data (i.e., analyses conducted on all patients who completed the study) and patients with imputed outcome data (i.e., imputed data from all withdrawal patients). The specifics of each assessment are presented in **Table 2** (22).

Analysis population	Definition
ITT	All randomised patients are included in efficacy analysis and adverse outcome analysis.
mITT	All randomised patients, except a defined patient group, are included in efficacy analysis and adverse outcome analysis.
As observed (AO)	Only patients who complete the trial are included in efficacy analysis and adverse outcome analysis.
Per protocol (PP)	Patients who adhered to terms of eligibility, interventions, and outcome assessment pre-specified in the protocol are included in efficacy analysis and adverse outcome analysis.
Other	None of the above mentioned categories fits the analysis population.
Unclear	It is unclear which analysis is applied.

Table 2: Analysis population for efficacy analysis and adverse outcome analysis (12;19;22).

The CONSORT statement has proposed equivalence between PP analysis and mITT covering exclusions of participants who did not adequately adhere to the protocol(57); we will honour the CONSORT statement and categorise analysis population as PP, if mITT is defined by exclusion of participants who did not adhere to the protocol.

When the number of patients who complete the follow-up examination is the same as those originally randomised, we will register the applied method of analysis as ITT, regardless of intended protocolised analysis population. Data analyses will assess both the *reported* and the *applied* method of analysis.

Trials conducting 'as observed' or 'per protocol' analysis will be excluded from subsequent analysis stages (8a-9b, figure 2) as they do not assess missing data or withdrawal patients, *per se*.

Missing data will be assessed from two perspectives. We will extract how the trials handle single sets of missing data noted as "BOCF,", "LOCF," "combined," "other," or "unclear." "Combined" will

refer to a combination of BOCF and LOCF, whereas "other" will refer to the use of other imputation techniques such as MI.

Subsequently we will determine whether trials distinguish among various withdrawal patients when handling missing data. We will register whether NRI (treatment fail) was applied to: a) patients who withdrew due to lack of efficacy; b) all patients who withdrew; or c) another defined patient group that withdrew but fit neither a or b.

When mITT is applied, we will assess the method of modification based on prior described categories (see **Figure 1** (12;19;22)). The modification will be registered as "Treatment" if final analysis comprised all randomised patients who had received at minimum one dose of the study drug. "Post-baseline assessment" will be noted as modification if final analysis included all patients who had at least one post-baseline assessment. The modification will be registered as "combined" if final analysis comprised all patients who had received at minimum one dose of the study drug and had at least one post-baseline assessment. Modifications that fail to fit into the three afore mentioned categories will be registered as "other." The number of modifications applied will be extracted as a numeric value.

A complete data extraction flowchart is presented in **Figure 2**. All data extraction will be done on trial level, except classification of intervention (step 3) and outcome assessment at primary endpoint (step 4) which will be extracted from each individual study arm.

Data extraction concerns both clinical and methodological relevant issues, some which may seem abundant to our study. We aim to use RHEUMATRIALS in future work concerning not only method but also clinical content specific work.

Evaluating methodological quality

The Cochrane risk of bias tool

The risk of bias within each trial will be assessed using the RoB tool as recommended by The Cochrane Collaboration (see below) (5;7). Each domain will be rated as "low," "high," or "unclear"

RoB. **Table 3** provides an overview of the Cochrane RoB components and their assessments, while **Table 4** provides an overview of other RoB components and their assessments. A domain will be rated as "unclear" if it fails to meet the criteria for "high" or "low" RoB.

To facilitate interpretation on the basis of the Cochrane risk of bias tool, each trial will subsequently be assigned an overall RoB. Overall RoB will be assessed based on the Cochrane bias components presented in **Table 3.** Overall RoB will be assessed tripartite as low risk (low for all Cochrane components), high risk (high for ≥1 Cochrane components), and unclear risk (unclear for ≥1 Cochrane components) (9). Overall RoB will also be assessed bipartite, categorized as low risk (low for all Cochrane components) or high risk (high or unclear for ≥1 Cochrane components) (3;58).

RoB item	Low RoB	High RoB
Sequence generation	It will be considered adequate if a random approach in the sequence generation process referred to a random number table, a random computer-generated number, coin tossing, drawing of lots, shuffling of cards, or throwing of dice. Multicentre trials described as randomised will be considered to have adequate sequence generation.	Date of birth, date of inclusion or admission, or record number of clinic/hospital is considered inadequate.
Allocation concealment	It will be considered adequate if there were no reasons to expect that the investigators responsible for inclusion were able to suspect which treatment was next. Both sequentially numbered, sealed, opaque envelopes and a central randomisation are considered adequate.	It will be regarded as inadequate if there is reason to expect that the investigators were able to suspect which treatment was next.
Blinding of patients, personnel, and outcome assessors	It will be considered adequate if the trials describe double-blinding.	It will be considered as inadequate if no blinding is described.
Incomplete outcome data	It will be considered adequate if missing data are distributed equally between intervention and control group. Further outcome data will be deemed adequate if data have been imputed using an appropriate technique and analyses based on the ITT population..	It will be considered inadequate if it is unclear how many patients are included in final analyses. Further, it is considered inadequate if no imputation technique is applied or if it is unclear how extensive the missing dataset is (i.e., unclear how many patients withdrew).
Selective reporting	It will be considered adequate if the chosen efficacy outcome (ACR20, ACR50 and/or ACR70) is reported in accordance with the usual contemporary RA protocols and reported at all specified time-points if more than one time point exists.	It will be considered inadequate if the chosen efficacy outcome (ACR20, ACR50 and/or ACR70) is not reported in accordance with the usual contemporary RA protocols, or is not reported at all specified time-points if more than one time point exists.

Table 3: *The Cochrane Risk of Bias tool.*

Other risk of bias components

Funding will be registered according to funding source, as described in **Figure 3** (12). Funding includes provision of manpower (authorship, statistical analysis, or other assistance), study materials (drug, placebo, assay kits, or similar materials), or grants (13). For-profit organisations will be defined as companies that might acquire financial gain or loss depending on the outcome of the trial. Trials partly financed by for-profit agencies will be registered as co-financed.

As an extension to the funding aspect, we will assess whether conflict of interest is reported as "none," denoted by "Yes" or "No."

To further assess methodological quality, we will note ("Yes" or "No") whether a flow-chart is publicly available.

Single or multicentre status will be determined through review (9). A trial will be considered a multicentre trial if more than one centre is involved. In case of missing information, trials will be classified as multicentre when there is reporting of both several ethics committees and different affiliations of authors. On the other hand, if the report stated only a single ethics committee and a single author affiliation, the trial will be classified as a single centre, unless other information calls for multicentre.

The geographical trial setting will be noted based on the continents participating in the trial: North America, Europe, South America, Asia, Africa and Australia. All the continents involved in the study will be registered.

RoB item	Low RoB	High RoB
Funding	No funding and not-for-profit funding will be considered as low RoB.	For-profit funding and co-financed funding will be considered high RoB. If funding is not reported, it will also be considered high RoB.
Conflict of interest	If conflict of interest is reported as "none," it will be considered low RoB.	If conflict of interest is not reported as "none," it will be considered high RoB.
Flowchart	If a flowchart is publicly available, it will be considered low RoB.	If a flowchart is not publicly available, it will be considered high RoB.
Number of study locations	It will be considered low RoB if more than one centre participated in the trial.	It will be considered high RoB if only one centre participated in the trial, or if it is unclear how many centres participated.

Table 4: *Risk of bias (RoB) components currently not included in the Cochrane RoB tool.*

Two reviewers will independently evaluate eligibility, risk of bias, assessment of ITT/mITT, and handling of missing data. Disagreements will be resolved by discussion until consensus.

Data synthesis

Data synthesis will assess ACR20 data at primary endpoint. We will present differences among different strata by estimating the ratio of odds ratios (ROR). We will assume that the observed number of responders in each arm of each trial has a binomial distribution. Accordingly, intervention effects will be modelled as log-odds ratios and outcomes will be coded so that ORs > 1 correspond to beneficial intervention effects. We will estimate the odds ratio (OR) of trials with the given characteristic using random-effects meta-analyses. For each drug corresponding to a meta-analysis, we will derive the difference between pooled estimates from trials with different trial characteristics (e.g., different ITT approaches). Formal tests of interaction between ITT analysis and estimated treatment benefits (active compared to control) will be performed using the following statistical methodology. After identifying all trials available for the different biological and targeted agents, we will record characteristics of individual studies (ITT: y/n; Type of ITT: ITT/mITT/AO/PP/Other/Unclear; Type of modification: Treatment/Post baseline assessment/Combined/Other; Handling of missing data in the trial: LOCF/BOCF/Other/Unclear; Handling of NRI: All wd/Lack Of Efficacy wd/Other wd/Unclear) and compare treatment effects within each biological or targeted agent. As characteristics will be assessed on trial level, analyses will assess comparison on trial level and not within trials.

We will fit *empirical Bayesian* hierarchical bias models using the generalized linear mixed models (GLMM) (59). Mean intervention effects may differ among trials with and without the reported study characteristic. Variation in bias among trials within biologics or targeted agents' trials is quantified and adjusted for with a fixed factor in the model. The GLMMs, like linear mixed models, assume normal (Gaussian) random effects. Conditional on these random effects, data can have any distribution in the exponential family. The exponential family comprises many of the elementary discrete and continuous distributions. The analyses will be performed using the GLIMMIX procedure in SAS (v. 9.2; SAS Institute Inc., Cary, NC, USA) (60). The syntax is similar to that of the MIXED procedure and includes CLASS, MODEL, and RANDOM statements. Using the GLIMMIX procedure, we will perform mixed-effects logistic regression with an arm-based, random-effects model within an *empirical Bayes* framework:

Proc Glimmix;
Class Trial Drug Group Characteristic;
Model Counts/Total = Group Characteristic Group*Characteristic / Solution;
Random Trial Trial*Group;
Lsmeans Group*Characteristic / cl ilink;
Run;

The PROC GLIMMIX statement invokes the procedure. The CLASS statement instructs the procedure to treat the variables Trial, Drug, Group, and Characteristic as classification variables. The MODEL statement specifies the response variable as a sample proportion using the r/N syntax: Counts/Total corresponds to Y_{iA}/N_{iA} for observations from Group A and to Y_{iB}/N_{iB} or observations from Group B. The SOLUTION option in the MODEL statement requests a listing of the fixed-effects parameter estimates. Because of the response/N syntax, the GLIMMIX procedure defaults to the binomial distribution, with the default logit link. The RANDOM statement specifies that the linear predictor contains intercept terms that randomly vary at the level of the Trial and Trial×Group effects. The default estimation technique in GLMMs is residual pseudo-likelihood (RSPL) with a subject-specific expansion. The default optimization technique for GLMMs is the Quasi-Newton method. Because a residual likelihood technique is used to compute the objective function, only the covariance parameters are participating in the optimization.

The LsMeans statement requests the least-squares means of the interaction between group (active vs. control) and the individual study characteristic effect on the logit scale. The CL option requests their confidence limits. The ILINK option adds estimates, standard errors, and confidence limits on the mean (probability) scale.

For secondary analyses we will consider possible interaction and confounding by other bias items presented in **Table 3 and 4**.

DISCUSSION

Biased results from RCTs ultimately put the patients at risk for being treated with pharmaceuticals with questionable efficacy and which may cause harm. Taking into account the expenses of accompanying RA treatment, this study is not only biomedical but also a socioeconomic necessity.

The term mITT is used to describe different methods for excluding participants post-randomised from analysis, thereby affecting and disregarding not only the ITT principle but also—and more importantly—the overriding purpose of ITT. Post-randomisation exclusions are known to induce bias, and theoretically mITT will introduce bias (23;24). Our study aim to establish if the bias is of practical concern, and focuses on the direction and magnitude of bias associated with mITT analyses. This study will present arguments as to why mITT approximates ITT or point to the problems concerning the use of mITT. As the term mITT embraces a broad notion of trials, we will delve into how the different types of modification influence effect size. This study may come out with neutral findings—which would not imply that overall bias associated with mITT analyses can be excluded, but may indicate that our study lacks the statistical power necessary to detect the bias. If some form of mITT can substitute ITT, guidelines regarding the use of mITT should be issued. In general this study examines many determinants, and therefore a risk of type I errors due to multiple comparisons exists and results must be interpreted carefully regardless of statistical significance (61).

This study is limited by the lack of agreement in how ITT and mITT are defined. Our mITT definition and categorisation is based on deviations described in the literature but have some shortcomings; e.g. in cases where only one post-baseline visit is required the mITT category post-baseline assessment will correspond to a completer's analysis.

As in other meta-epidemiological studies, we are limited by the many sources of heterogeneity, e.g. differences in disease duration, type of RA population and intervention dose. As meta-epidemiological studies concerns methodology and do not aim at establishing the empirical evidence for an intervention effect, this underlying premise of heterogeneity can be viewed as acceptable. However, heterogeneity should always be borne in mind when interpreting results.

Our primary objective is to examine whether mITT is associated with different effect sizes, implying empirical evidence for bias in treatment effects. ITT prevents attrition bias when evaluating treatment assignment but may not provide a true estimate of treatment effect if some patients are non-adherent (62). As the term "bias" comprises deviation from the true intervention effect, it can be perceived as misleading to regard systematic errors in treatment effect between mITT and ITT analyses as "bias," given that ITT analysis may fail to provide a true evaluation of the intervention effect. However, ITT analysis is recommended as the least biased way to estimate intervention effects (7) and concerns regarding the systematic errors between mITT and ITT remain, regardless of terminology. This project builds on the premise that the trials included are otherwise less prone to bias, although there is no guarantee that recent trials on biologics and targeted interventions will be at low risk of bias.

This study may point to potential bias and disadvantages in the handling of missing data in RCTs, otherwise known for having a low risk of bias compared with other study designs (63). SI has been criticised on a theoretical level, but its implication on efficacy outcomes in RA trials is uncharted. Accordingly, this study may provide empirical evidence that can support or contradict existing critics. Regardless of our findings one should always be careful when interpreting results from trials where data are missing and consider the reasons for missing data and potential impact on effect estimates(7;64).

The study examines potential bias associated with industry funding. It may prove difficult to assess bias, as most RA trials concerning biological and targeted agents have some degree of industry input. Being unable to reject industry bias and unable to estimate the influence, direction, and magnitude of such, the validity of trial results in this industry-permeated field of research is open to conjecture (14).

Dissemination

First author Anna Dossing will draft a paper describing the systematic review; the meta-epidemiological study will be disseminated by peer-review publication and conference presentations.

HISTORY

Protocol first published: 11. December 2013

CONTRIBUTION OF AUTHORS

All authors fulfil the ICMJE guidelines for authorship. AD, ST, DEF, CG, JB, BBH, HB, RC participated in the conception and design of this protocol. RC provided statistical advice for the design and analysis. AD drafted the manuscript. AD, ST, DEF, CG, JB, BBH, HB, RC critically reviewed the manuscript for important intellectual content and approved the final version.

FUNDING

This research received grants from the Michaelsen foundation. No sponsor was involved in study design, and no sponsor will have authority in the collection, management, analysis, and interpretation of data. Writing of the report and the decision to submit the results for publication is strictly made by the authors. Musculoskeletal Statistics Unit, The Parker Institute, is supported by grants from the Oak Foundation. The Copenhagen Trial Unit is funded by the Danish state.

DECLARATION OF INTEREST

The authors declare no conflicts of interest.

ABBREVIATIONS

ACR   American College of Rheumatology

AO   As observed

bDMARDs Biological disease-modifying anti-rheumatic drugs

BOCF	Baseline Observation Carried Forward
CRP	C-reactive protein
csDMARDs	Conventional synthetic disease-modifying anti-rheumatic drugs
DAS28	European Disease Activity Score
DMARDs	Disease-modifying anti-rheumatic drugs
EMA	European Medicines Agency
ESR	Erythrocyt sedimentation rate
EULAR	European League Against Rheumatism
FDA	US Food and Drug Administration
GCP	Good clinical practice
GLMM	Generalized linear mixed models
ICTRP	WHO Clinical trials Portal
IR	Inadequate response
ITT	Intention-To-Treat
LOCF	Last Observation Carried Forward
MAR	Missing at random
MCAR	Missing completely at random
MI	Multiple imputation
mITT	modified Intention-To-Treat
MNAR	Missing not at random
MTX	Methotrexat
NRI	Non-responder imputation
OR	Odds ratio
PP	Per protocol
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RA	Rheumatoid Arthritis
RCT	Randomised controlled trial
RF	Rheumatoid factor
RoB	Risk of bias
ROR	Ratio of odds ratios

RSPL	Residual pseudo-likelihood
sDMARDs	Synthetic disease-modifying anti-rheumatic drugs
SI	Single imputation
tsDMARD	Targeted synthetic disease-modifying anti-rheumatic drugs
wd	Withdrawal patients

REFERENCES

(1) Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials.   BMJ 2001 Jul 7;323(7303):42-6.

(2) Sterne JA, Juni P, Schulz KF, et al. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. Stat Med 2002 Jun 15;21(11):1513-24.

(3) Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008 Mar 15;336(7644):601-5.

(4) Savovic J, Jones H, Altman D, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. Health Technol Assess 2012 Sep;16(35):1-82.

(5) Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

(6) Gluud LL. Bias in clinical intervention research. Am J Epidemiol 2006 Mar 15;163(6):493-501.

(7) Higgins J, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 ed. The Cochrane Collaboration; 2011.

(8) Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med 2001 Dec 4;135(11):982-9.

(9) Dechartres A, Boutron I, Trinquart L, et al. Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study. Ann Intern Med 2011 Jul 5;155(1):39-51.

(10) Bafeta A, Dechartres A, Trinquart L, et al. Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. BMJ 2012;344:e813.

(11) Dechartres A, Trinquart L, Boutron I, et al. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ 2013;346:f2304.

(12) Montedori A, Bonacini MI, Casazza G, et al. Modified versus standard intention-to-treat reporting: are there differences in methodological quality, sponsorship, and findings in randomized trials? A cross-sectional study. Trials 2011;12:58.

(13) Als-Nielsen B, Chen W, Gluud C, et al. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 2003 Aug 20;290(7):921-8.

(14) Lundh A, Sismondo S, Lexchin J, et al. Industry sponsorship and research outcome. Cochrane Database Syst Rev 2012;12:MR000033.

(15) Bero LA. Why the Cochrane risk of bias tool should include funding source as a standard item. Cochrane Database Syst Rev 2013;12:ED000075.

(16) Sterne JA. Why the Cochrane risk of bias tool should not include funding source as a standard item. Cochrane Database Syst Rev 2013;12:ED000076.

(17) Shrier I, Steele RJ, Verhagen E, et al. Beyond intention to treat: what is the right question? Clin Trials 2014 Feb;11(1):28-37.

(18) Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say? Clin Trials 2007;4(4):350-6.

(19) Kruse RL, Alper BS, Reust C, et al. Intention-to-treat analysis: who is in? Who is out? J Fam Pract 2002 Nov;51(11):969-71.

(20) Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999 Sep 11;319(7211):670-4.

(21) Baron G, Boutron I, Giraudeau B, et al. Violation of the intent-to-treat principle and rate of missing data in superiority trials assessing structural outcomes in rheumatic diseases. Arthritis Rheum 2005 Jun;52(6):1858-65.

(22) Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: systematic review. BMJ 2010;340:c2697.

(23) Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. Int J Epidemiol 2005 Feb;34(1):79-87.

(24) Nuesch 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.

(25) Altman DG, Bland JM. Missing data. BMJ 2007 Feb 24;334(7590):424.

(26) Boers M. Missing data in trials: do we have to keep carrying the last observation forward? Arthritis Rheum 2008 Jan 15;59(1):2-3.

(27) EMA. Guideline on missing data in confirmatory clinical trials. London: European Medicines Agency (EMA); 2010.

(28) White IR, Horton NJ, Carpenter J, et al. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ 2011;342:d40.

(29) Shao J, Jordan DC, Pritchett YL. Baseline observation carry forward: reasoning, properties, and practical issues. J Biopharm Stat 2009 Jul;19(4):672-84.

(30) Molnar FJ, Hutton B, Fergusson D. Does analysis using "last observation carried forward" introduce bias in dementia research? CMAJ 2008 Oct 7;179(8):751-3.

(31) Altman DG. Missing outcomes in randomized trials: addressing the dilemma. Open Med 2009;3(2):e51-e53.

(32) Baron G, Boutron I, Giraudeau B, et al. Reporting of radiographic methods in randomised controlled trials assessing structural outcomes in rheumatoid arthritis. Ann Rheum Dis 2007 May;66(5):651-7.

(33) Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.

(34) FDA. Guidance for Industry, Clinical Development Programs for Drugs, Devices, and Biological Products for the Treatment of Rheumatoid Arthritis (RA). 1999.

(35) Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988 Mar;31(3):315-24.

(36) Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995 Jan;38(1):44-8.

(37) Nielen MM, van SD, Reesink HW, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004 Feb;50(2):380-6.

(38) Dahlqvist SR. Rheumatoid arthritis increased the risk for myocardial infarction in women. ACP J Club 2003 Sep;139(2):50.

(39) Scott DL, Symmons DP, Coulton BL, et al. Long-term outcome of treating rheumatoid arthritis: results after 20 years. Lancet 1987 May 16;1(8542):1108-11.

(40) Welsing PM, van Gestel AM, Swinkels HL, et al. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum 2001 Sep;44(9):2009-17.

(41) van DH, van AJ, Lard LR, et al. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. Arthritis Rheum 2007 May;56(5):1424-32.

For peer review only

(42) Smolen JS, Landewe R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. Ann Rheum Dis 2010 Jun;69(6):964-75.

(43) Furst DE, Keystone EC, Braun J, et al. Updated consensus statement on biological agents for the treatment of rheumatic diseases, 2011. Ann Rheum Dis 2012 Apr;71 Suppl 2:i2-45.

(44) Tugwell P, Singh JA, Wells GA. Biologicals for rheumatoid arthritis. BMJ 2011;343:d4027.

(45) Smolen JS, Landewe 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 2013 Oct 25.

(46) Strand V, Sokolove J. Randomized controlled trial design in rheumatoid arthritis: the past decade. Arthritis Res Ther 2009;11(1):205.

(47) Stein CM, Pincus T. Placebo-controlled studies in rheumatoid arthritis: ethical issues. Lancet 1999 Jan 30;353(9150):400-3.

(48) EMA. Guideline on clinical investigation of medicinal products other than NSAIDs for treatment of rheumatoid arthritis. 2011.

(49) FDA. Guidance for Indusry, Rheumatoid Arthritis: Developing Drug Products for Treatment. 2013.

(50) Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995 Jun;38(6):727-35.

(51) EMA. Points to consider on clinical investigation of medicinal products other than NSAIDS for treatment of rheumatoid arthritis. London: European Medicines Agency (EMA); 2003.

(52) EMA. Draft guideline on clinical investigation of medicinal products other than non-steroidal anti-inflammatory drugs for treatment of rheumatoid arthritis. London: European Medicines Agency (EMA); 2011.

(53) Aletaha D, Landewe R, Karonitsch T, et al.Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Arthritis Rheum 2008 Oct 15;59(10):1371-7.

(54) Dossing A, Tarp S, Furst DE, et al. Attrition bias in rheumatoid arthritis randomised trials with different modified intention-to-treat approaches: a meta-epidemiological study. PROSPERO 2013 Dec 11;CRD42013006702.

(55) Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med 2009 Aug 18;151(4):W65-W94.

(56) Wells G, Becker JC, Teng J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009 Jun;68(6):954-60.

(57) 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(1):28-55.

(58) Savovic J, Jones HE, Altman DG, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med 2012 Sep 18;157(6):429-38.

(59) Platt RW, Leroux BG, Breslow N. Generalized linear mixed models for meta-analysis. Stat Med 1999 Mar 30;18(6):643-54.

(60) Singh JA, Christensen R, Wells GA, et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. Cochrane Database Syst Rev 2009;(4):CD007848.

(61) Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004 Jun 15;23(11):1663-82.

(62) Shrier I, Steele RJ, Verhagen E, et al. Beyond intention to treat: What is the right question? Clin Trials 2013 Oct 3.

(63) Jakobsen J C & Gluud C. The Necessity of Randomized Clinical Trials. British Journal og Medicine and Clinical Research 2013 May 1;3(4):1453-68.

(64) Akl EA, Briel M, You JJ, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. BMJ 2012;344:e2809.

FIGURE LEGENDS

Figure 1: Overview of modified intention-to-treat (mITT) categories.

The four categories are based on the most common deviations described in the literature.

Figure 2: Data extraction flowchart.

csDMARD: Conventional synthetic disease-modifying anti-rheumatic drugs. bDMARD: Biological

disease-modifying anti-rheumatic drugs. tsDMARD: Targeted synthetic disease-modifying anti-

rheumatic drugs. IR: Inadequate response. ACR20, 50, 70: American College of Rheumatology 20%,

50%, 70% improvement in disease activity respectively. DAS28: European Disease Activity Score.

ITT: Intention-To-Treat. mITT: modified Intention-To-Treat. AO: As observed. PP: Per protocol.

LOCF: Last Observation Carried Forward. BOCF: Baseline Observation Carried Forward. NRI: Non-

responder imputation. Wd: Withdrawal patients

Figure 3: Funding sources.

Interpreting trial results following use of different intention-to-treat approaches for preventing attrition bias: A meta-epidemiological study protocol

Anna Dossing[1,2], Simon Tarp[1], Daniel E. Furst[3], Christian Gluud[4], Joseph Beyene[5], Bjarke B. Hansen[1], Henning Bliddal[1], Robin Christensen[1]

Affiliations
1: Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospitals, Bispebjerg and Frederiksberg, Denmark
2: Department of Clinical Medicine, Faculty of Medical and Health Sciences, University of Copenhagen, Denmark
3: University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, USA
4: Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
5: Department of Clinical Epidemiology & Biostatistics, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

Correspondence to:
Dr. Robin Christensen, MSc, PhD
Head of Musculoskeletal Statistics Unit;
The Parker Institute, Department of Rheumatology.
Copenhagen University Hospital, Bispebjerg and Frederiksberg.
Nordre Fasanvej 57
DK-2000 Copenhagen Frederiksberg
Denmark
Phone: +45 3816 4165
Fax: +45 3816 4159
E-Mail: Robin.Christensen@frh.regionh.dk

Word count: (excluding title page, abstract, references, figures and tables): 4,654

ABSTRACT

Introduction: When subjects drop out of randomised clinical trials, as frequently happens, the intention-to-treat principle do~~es~~ not apply, potentially leading to attrition bias. Data lost from patient dropout/lack of follow-up are statistically addressed by imputing, a procedure prone to bias. Deviations from the original definition of ITT ~~have led to various terminologies, such as~~are referred to as modified ITT (mITT). As yet, the impact of the potential bias associated with mITT has not been assessed. Our objective is to investigate potential bias and disadvantages of performing mITT and evaluate possible concerns when executing different mITT approaches in meta-analyses.

Methods and analysis: Using meta-epidemiology on randomised trials considered less prone to bias (i.e., good internal validity) and assessing biological or targeted agents in patients with rheumatoid arthritis (RA), we will meta-analyse data from 10 biological and targeted drugs based on collections of trials that would correspond to 10 individual meta-analyses.

Ethics and dissemination: This study will enhance transparency for evaluating mITT treatment effects described in meta-analyses. The intended audience will include health care researchers, policymakers and clinicians. Results of the study will be disseminated by peer-review publication.

Protocol registration: In ~~prospero~~ PROSPERO CRD42013006702, 11. December 2013

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This meta-epidemiological study is the first to focus on bias associated with mITT and its impact on effect size.

- This study will influence how results from RCTs and meta-analysis should be interpreted; this study will provide a framework for designing, conducting and reporting future RCTs in order to minimise attrition bias.

- In practice, various definitions of both ITT and mITT are used. Some may argue that our definitions are not stringent enough or inverse overly stringent.

INTRODUCTION

Meta-epidemiological research

Inadequate quality of trials may distort the results from meta-analyses and systematic reviews (1). Consequently, meta-epidemiological studies are carried out to quantitatively evaluate bias across many randomised clinical trials (RCTs) in different meta-analyses (2-4). Overall flaws in the design, conduct, analysis, and reporting of RCTs can introduce bias—systematic errors that lead to overestimating or underestimating the benefits and the harms of treatment (5;6).

Methodological quality and risk of bias assessment

The Cochrane Collaboration's Handbook for Systematic Reviews of Interventions provides guidance to authors to critically review trial outcome using the risk of bias (RoB) assessment tool (5). The RoB tool requires authors to evaluate the well-established strengths and limitations of RCTs, including sequence generation, allocation concealment, blinding of participants, personnel and outcome assessors, loss to follow-up with failure to apply the intention-to-treat (ITT) principle, and selective outcome reporting (5;7). In the course of meta-epidemiological studies, other sources of bias in RCTs have been identified, such as significant discrepancies favouring intervention in single (rather than multicentre) trials, in trials with small (rather than large) sample sizes, and in using subjective (rather than objective) outcome measures (3;8-11). Most recently, funding source has become a distinct possibility as a source of bias, with for-profit organisation funding likely

favouring pro-intervention results (12-14), however there is an on-going debate as to whether funding should be regarded as a RoB item (15;16).

The intention-to-treat principle

The ITT principle has two main rules: 1) it requires all participants in a RCT to be analysed according to their original allocation, regardless of their adherence to the trial protocol; and 2) all randomised participants must be included in the analysis. The first rule acts to conserve randomisation, which is executed to avoid selection bias and thereby produces treatment groups in which the distributions of prognostic factors—both known and unknown—are similar. The second rule serves to avoid attrition bias when evaluating a treatment assignment (Shrier, 2013 1928 /id)(17). Attrition bias is attributed to systematic differences between groups in withdrawals from the study (7).

Frequently, not all participants in RCTs are analysed as they were initially randomised (18-21). Various deviations from the definition of ITT have led to the vaguely defined term "modified ITT" (mITT), which may further compromise true randomisation. The incidence of trials reporting the use of mITT has increased over the years (22). Trials that report mITT exclude patients from analysis post-randomisation compared to trials reporting on the ITT population (12). Although post-randomisation exclusion is known to produce bias (23;24), the potential magnitude and direction of bias associated with mITT is unknown.

Based on the most common deviations from ITT described in the literature (19;20;22), mITT can be divided into four categories, as illustrated in **Figure 1**.

	Treatment	All patients who have received minimum 1 dose of study drug/placebo.
	Post-baseline assessment	All patients who have had at least 1 post-baseline assessment.
mITT	Combined	All patients who had received minimum 1 dose of study drug and had at least 1 post-baseline assessment.
	Other	Exclusion of patients who fit none of the categories above.

Figure 1: Overview of modified intention-to-treat (mITT) categories. The four categories are based on the most common deviations described in the literature.

Handling of missing data

Executing ITT ~~and mITT~~ analyse~~i~~s can be difficult, as missing data is common (25;26). Missing data also affects mITT analysis unless mITT is defined by a complete case analysis. Missing data comprises single missing data points as well as missing datasets due to withdrawal (26). Data can be missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). MNAR occurs when missing data depend on unobserved values: and the missing data may lead to bias (25-27). Regardless of trial design and objective, it is likely that some of the data will be MNAR ~~to a degree~~.

Problems from missing data in RCTs can be somewhat mitigated by data imputation—a procedure by which missing data are replaced with a statistically founded ~~presumably conservative~~ estimate (27;28). Two single imputation (SI) methods often used are "*last observation carried forward"* (LOCF) or "*baseline observation carried forward"* (BOCF). LOCF uses the last observed value in place of the missing outcome, whereas BOCF uses the baseline observation as the value replacing the outcome (27;29). LOCF and BOCF have been widely criticised~~, as they are based on the assumption that lack of treatment equates with a halt in disease progression~~ (21;30;31) and the use of SI risk introducing bias, e.g. by minimising variation (27;32). A newer and more promising imputation method is multiple imputation (MI), which involves creating several different plausible imputed datasets based on a Bayesian approach and then combining the results obtained from each of them, hence reducing the risk of underestimating variance (27;32;33).

Loss to follow-up is frequent in RCTs and can be attributed to a variety of causes (23;25). Most patients drop out of trials due to lack of efficacy, adverse effects of treatment, or both (26). When handling a dichotomous outcome in a responder analysis, it is often considered suitable to attribute patients' withdrawal to lack of efficacy, and therefore assume treatment failure (27;34) (also referred to as non-responder imputation (NRI)). NRI is applied inconsistently to different withdrawal populations, e.g. all patients who withdraw are considered treatment failures as opposed to only patients who withdraw due to lack of efficacy are considered treatment failures. In principle there should be no objection to applying NRI to different withdrawal populations (27) and no empirical evidence have documented whether it affects trial results.

Rationale for this meta-epidemiological study

Our extensive search did not find any previous systematic assessment of empirical evidence for bias associated with mITT (23;24). Furthermore, mITT has not been assessed as to whether the type or number of modifications applied affects the estimated efficacy outcome differently. To investigate potential bias, we focus on trials assessing biological (or targeted) interventions in patients with rheumatoid arthritis (RA), as these trials are relatively recent and we therefore anticipate that these yield reasonable internal validity. With regards to imputation, we wish to examine how imputation of missing data affects RA trials investigating clinical response, as this have, to our knowledge, not previously been done. The study also aims to shed a light on the effect of selectively applying NRI to different withdrawal populations, which so far has remained unexplored.

Biological and targeted agents for rheumatoid arthritis

RA is a chronic inflammatory autoimmune disease characterised by joint swelling and joint tenderness with destruction of synovial joints primarily affecting the hands (35-38). The inflammatory load drives the destructive progression of the disease and leads to severe disability and premature mortality (39;40). Diagnosis of RA requires that patients have a minimum of four criteria that persist for 6 months (35). The seven qualifying criteria are: 1) morning stiffness; 2)

arthritis of three or more joints or joint areas; 3) arthritis of hand joints; 4) symmetric arthritis; 5) rheumatoid nodules; 6) serum rheumatoid factor (RF); and 7) radiographic changes.

Treatment of RA encompasses multiple interventions. Though RA has been shown to progress over time despite treatment (39), early therapeutic intervention improves clinical outcomes and reduces the accrual of joint damage and disability (41). RA can be managed using disease-modifying anti-rheumatic drugs (DMARDs). DMARDs form two major classes: synthetic chemical compounds (sDMARDs) and biological agents (bDMARDs). sDMARDs comprise the conventional DMARDs (csDMARDs) (e.g., methotrexat [MTX] and sulfasalazin) and the new targeted sDMARDs (tsDMARDs); e.g., tofacitinib). bDMARDs is a heterogeneous group of pharmaceuticals including abatacept, adalimumab, anakinra certolizumab, etanercept, golimumab, infliximab, rituximab, and tocilizumab that help to control the autoimmune inflammation associated with arthritis (42-45). The DMARDs of interest are summarised in **Table 1** (45).

DMARDs	Pharmaceutical
csDMARDs	methotrexat, sulfosalazin, leflunomid, hydroxychloroquine
tsDMARD	Tofacitinib
bDMARDs	adalimumab, abatacept, anakinra, certolizumab pegol, etanercept, golimumab, infliximab, rituximab, tocilizumab

Table 1: Overview of disease-modifying anti-rheumatic drugs (DMARD) groups: conventional DMARDs (csDMARDs); the new targeted sDMARDs (tsDMARDs); and biological agents (bDMARDs).

In RA RCTs on biological and targeted agents, the control group is commonly treated with MTX, as lack of treatment can lead to irreversible loss of physical function in RA patients (46;47). Rescue therapy (e.g., regulation of dose or addition of MTX or bDMARD or tsDMARD) is acknowledged by both the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA) and is typically offered when treatment response is evaluated as inadequate (48;49). Because patients who receive rescue therapy are commonly encountered as withdrawal patients in the trials protocols; rescue therapy challenges the ITT principle and may contribute to attrition bias.

The treatment effect in RA RCTs is typically evaluated using the American College of Rheumatology (ACR) response criteria and/or variations of the European Disease Activity Score (DAS28) (46). The DAS28 score discriminates between high and low disease activity (36). The ACR response criteria's definition of improvement in RA trials includes improvement in the joint counts and improvement in at least three of the five following: 1) patient assessment; 2) physician assessment; 3) erythrocyte sedimentation rate; 4) pain scale; and 5) functional questionnaire (50): Improvement is described as ACR20, ACR50 or ACR70 and refers to a 20%, 50%, or 70% improvement in tender and swollen joint count and 20%, 50%, or 70% improvement in three of the five core set measures listed above. The EMA, the European League Against Rheumatism (EULAR), and the ACR consent that validated composite clinical outcomes such as the ACR response criteria should be used to document efficacy of treatment. Specifically the ACR20, ACR50, or ACR70 should be used to document signs and symptoms after 3–6 months, referring respectively to a 20%, 50%, or 70% improvement in disease activity (51-53).

Objective

Our primary objective is to examine

1. Whether mITT is associated with different effect sizes compared to ITT.

Secondarily we wish to examine

2a. How the choice of imputation technique influences the effect size.

2b. How different ITT populations affect effect size.

Hypotheses

We hypothesise that different ITT populations are associated with differences in treatment effect sizes. We expect the level of statistical significance will depend on the type of modification, the number of modifications applied, and the percentage of patients excluded from final efficacy analysis. Furthermore, we hypothesise that the use of different mITT approaches is strongly associated with industry funding (12) and these modifications are applied to artificially increase the effect size for the apparent benefit in trials directly related to a specific pharmaceutical company.

METHODS AND ANALYSIS

Protocol and registration

Our protocol is registered on PROSPERO (CRD42013006702)(54); our manuscript conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for reporting systematic reviews and meta-analyses (55).

Eligibility criteria

All RA RCTs concerning EMA- or FDA-approved biologics and targeted agents evaluating efficacy as ACR20, ACR50, or ACR70 will be considered eligible, independent of whether they include ITT or mITT analysis. We will include published RCTs. Open-label studies will be excluded from our analysis because performance and detection bias are inherent when the outcome is not considered an objective endpoint (3).

Search and selection of trials and meta-analyses

We will search PubMed, EMBASE, The Cochrane Central Register of Controlled Trials (CENTRAL), and LILACS using a combination of keywords and text words related to RA. The PubMed search strategy is listed below. See **Appendix 1** for the EMBASE, the Cochrane Library, and LILACS search strategies. The World Health Organisation (WHO) Clinical trials Portal (ICTRP), clinicaltrials.gov, FDA, EMA and pharmaceutical companies' trial result databases will be searched to identify unpublished data.

PubMed search strategy: (((("Receptors, Tumor Necrosis Factor"[nm] OR TNFR:Fc OR "TNFR-Fc fusion protein"[Supplementary Concept] OR "TNFR-Fc fusion protein"[All Fields] OR "etanercept"[All Fields] OR "enbrel"[All Fields]) OR ("infliximab"[Supplementary Concept] OR "infliximab"[All Fields] OR "remicade"[All Fields] OR "mab ca2"[All Fields] OR "monoclonal antibody ca2"[All Fields]) OR ("adalimumab"[Supplementary Concept] OR "adalimumab"[All Fields] OR "humira"[All Fields]) OR ("interleukin 1 receptor antagonist protein"[MeSH Terms] OR "interleukin 1 receptor antagonist protein"[All Fields] OR "anakinra"[All Fields] OR "kineret"[All

Fields] OR "antril"[All Fields]) OR ("abatacept"[Supplementary Concept] OR "abatacept"[All Fields])

OR CTLA4Ig[All Fields] OR "orencia"[All Fields]) OR ("rituximab"[Supplementary Concept] OR

"rituximab"[All Fields] OR "rituxan"[All Fields] OR "idec c2b8"[All Fields]) OR ("golimumab"[All

Fields] OR "golimumab"[Supplementary Concept] OR "simponi"[All Fields] OR "cnto-148"[All

Fields] OR ("cnto"[All Fields] AND "148"[All Fields])) OR ("tocilizumab"[All Fields] OR

"tocilizumab"[Supplementary Concept] OR "atlizumab"[All Fields] OR "actemra"[All Fields]) OR

("certolizumab"[All Fields] OR "certolizumab pegol"[Supplementary Concept] OR "CDP870"[All

Fields] OR ("cdp"[All Fields] AND "870"[All Fields]) OR "cimzia"[All Fields]) OR

("tofacitinib"[Supplementary Concept] OR "tofacitinib"[All Fields]) OR ("Antibodies,

Monoclonal"[Mesh] OR "Monokines"[Mesh] OR "Receptors, Interleukin-1"[Mesh] OR "Receptors,

Interleukin-6"[Mesh])) AND ("Randomized Controlled Trial"[ptyp] OR "Controlled Clinical

Trial"[ptyp] OR "Multicenter Study"[ptyp] OR "randomized"[tiab] OR "randomised"[tiab] OR

"placebo"[tiab] OR "randomly"[tiab] OR "trial"[tiab] OR randomized controlled trials[mh] OR

random allocation[mh] OR double-blind method[mh] OR single-blind method[mh]) AND ("Arthritis,

Rheumatoid"[MeSH Terms] OR (Rheumatoid[text word] AND arthriti*[text word]))) NOT

(animals[mh] NOT human[mh]).

Date for final database update will be stated.

Data extraction

All RCT publications will be assigned an ID, and we will extract information on author, year of

publication, journal of publication, company study name, company study number, and registration

number. All characteristics will be typed into a custom made database, (RHEUMATRIALS). We will

extract baseline characteristics and inclusion criteria. RA will be classified on the basis of mean

disease duration as early (≤ 6 months), established (6 months to 2 years), or late (> 2 years). Study

duration until registration of primary outcome and duration of the longest placebo-controlled

period will be extracted. We will extract type of primary outcomes (i.e., safety, efficacy, or both)

and the name of the primary outcome (e.g., ACR20).

Trials will be classified based on the included patient populations. We will distinguish among three possibilities: whether patients have exhausted all csDMARD treatment options, whether they have had inadequate response to csDMARD treatment and are candidates for biologics or a targeted agent, or whether they have experienced an inadequate response to a bDMARD or tsDMARD.

Regarding the intervention, we will extract the main drug of interest, the allocation, the form of administration, dosing of the drug, and registration of whether dose is equivalent to standard dose. We will register if treatment with csDMARD concomitant to the intervention was allowed. Offering of rescue will be registered as "yes" or "no": time-point and the content of rescue will also be registered.

We will extract the total number of patients reaching the primary outcome and the total number of patients in the analysis. For efficacy outcomes, we will assess ACR20, ACR50, and ACR70. For sensitivity ~~When available~~, we will extract the continuous efficacy outcome DAS28 as the mean change from baseline in each intervention group with the corresponding dispersion, when available. If DAS28 was reported based on C-reactive protein (CRP) and on erythrocyt sedimentation rate (ESR), we will extract only DAS28 (ESR), as both outcomes are in aggreement (56). For safety outcomes, we will assess overall withdrawals, withdrawals due to adverse events, and number of serious adverse events.

Intention-to-treat analysis on efficacy outcome

We will extract the number of patients randomised in the trial in total, the number of randomised patients per intervention group, the number of patients completing each intervention group, and the number of patients included in efficacy analysis per intervention group. The method of efficacy analysis will be extracted on two levels. The first level will assess the *reported* protocolised method of analysis. The second level will register the *applied* method of analysis. The analysis population will be categorised as: "ITT," "mITT," "as observed," "per protocol," "other," or "unclear." Categorisation will be based on the population included in ACR20 efficacy analysis. A full ITT analysis will comprise both patients with clinically assessed outcome data (i.e., analyses conducted on all patients who completed the study) and patients with imputed outcome data (i.e.,

imputed data from all withdrawal patients). The specifics of each assessment are presented in

Table 2 (22).

Analysis population	Definition
ITT	All randomised patients are included in efficacy analysis and adverse outcome analysis.
mITT	All randomised patients, except a defined patient group, are included in efficacy analysis and adverse outcome analysis.
As observed (AO)	Only patients who complete the trial are included in efficacy analysis and adverse outcome analysis.
Per protocol (PP)	Patients who adhered to terms of eligibility, interventions, and outcome assessment pre-specified in the protocol are included in efficacy analysis and adverse outcome analysis.
Other	None of the above mentioned categories fits the analysis population.
Unclear	It is unclear which analysis is applied.

Table 2: *Analysis population for efficacy analysis and adverse outcome analysis (12;19;22).*

The CONSORT statement has proposed equivalence between PP analysis and mITT covering exclusions of participants who did not adequately adhere to the protocol(57); we will honour the CONSORT statement and categorise analysis population as PP, if mITT is defined by exclusion of participants who did not adhere to the protocol.

When the number of patients who complete the follow-up examination is the same as those originally randomised, we will register the applied method of analysis as ITT, regardless of intended protocolised analysis population. Data analyses will assess both the *reported* and the *applied* method of analysis.

Trials conducting 'as observed' or 'per protocol' analysis will be excluded from subsequent analysis stages (8a-9b, figure 2) as they do not assess missing data or withdrawal patients, *per se*.

Missing data will be assessed from two perspectives. We will extract how the trials handle single sets of missing data noted as "BOCF,", "LOCF," "combined," "other," or "unclear." "Combined" will refer to a combination of BOCF and LOCF, whereas "other" will refer to the use of other imputation techniques such as MI.

Subsequently we will determine whether trials distinguish among various withdrawal patients when handling missing data. We will register whether NRI (treatment fail) was applied to: a) patients who withdrew due to lack of efficacy; b) all patients who withdrew; or c) another defined patient group that withdrew but fit neither a or b.

When mITT is applied, we will assess the method of modification based on prior described categories (see **Figure 1** (12;19;22)). The modification will be registered as "Treatment" if final analysis comprised all randomised patients who had received at minimum one dose of the study drug. "Post-baseline assessment" will be noted as modification if final analysis included all patients who had at least one post-baseline assessment. The modification will be registered as "combined" if final analysis comprised all patients who had received at minimum one dose of the study drug and had at least one post-baseline assessment. Modifications that fail to fit into the three afore mentioned categories will be registered as "other." The number of modifications applied will be extracted as a numeric value.

A complete data extraction flowchart is presented in **Figure 2**. All data extraction will be done on trial level, except classification of intervention (step 3) and outcome assessment at primary endpoint (step 4) which will be extracted from each individual study arm.

Data extraction concerns both clinical and methodological relevant issues, some which may seem abundant to our study. We aim to use RHEUMATRIALS in future work concerning not only method but also clinical content specific work.

1. Trial background information
- Author
- Year and journal of publication
- Company study name and number
- Registration number
- Primary outcome

2. Classification of population

Classification of prior treatment and prior treatment responses: csDMARD naive, csDMARD inadequate response (IR), bDMARD or tsDMARD IR.

3. Classification of intervention
- Intervention drug
- Dose
- Form of administration
- Frequency of administration
- Concomitant csDMARD treatment

4. Outcome assessment at primary endpoint

Efficacy: ACR20, 50, and 70 will be assessed. DAS28 data will be extracted as well.
Safety: Overall withdrawal, withdrawal due to adverse events, and number of serious adverse events will be assessed.
We will assess the total number of events and the total number of patients included for analysis.

5. Reported assessment of efficacy analyses

Reported population for efficacy analysis

ITT   mITT   AO   PP   Other   Unclear

See table 2 for elaboration.

6. Reported method of analysis applied?   Yes/No

7. Applied assessment of efficacy analyses

Applied population for efficacy analysis

ITT   mITT   AO   PP   Other   Unclear

See table 2 for elaboration.

Objective 1

8a. Handling of missing data points

Imputation of single missing data points

LOCF   BOCF   Other   Unclear

9a. Handling of withdrawal patients (wd)

When was NRI applied?

All wd   Lack Of Efficacy wd   Other wd   Unclear

Objective 2a

8b. Type of modification applied

mITT

Treatment   Post-baseline assessment   Combined   Other

See figure 1 for elaboration.

9b. No. of modifications applied   Numeric value

Objective 2b

Figure 2: *Data extraction flowchart.*

Evaluating methodological quality

The Cochrane risk of bias tool

The risk of bias within each trial will be assessed using the RoB tool as recommended by The Cochrane Collaboration (see below) (5;7). Each domain will be rated as "low," "high," or "unclear" RoB. **Table 3** provides an overview of the Cochrane RoB components and their assessments, while **Table 4** provides an overview of other RoB components and their assessments. A domain will be rated as "unclear" if it fails to meet the criteria for "high" or "low" RoB.

To facilitate interpretation on the basis of the Cochrane risk of bias tool, each trial will subsequently be assigned an overall RoB. Overall RoB will be assessed based on the Cochrane bias components presented in **Table 3.** Overall RoB will be assessed tripartite as low risk (low for all Cochrane componentskey domains), high risk (high for ≥1 Cochrane componentskey domains), and unclear risk (unclear for ≥1 Cochrane componentskey domains) (9). Overall RoB will also be assessed bipartite, categorized as low risk (low for all Cochrane componentskey domains) or high risk (high or unclear for ≥1 Cochrane componentskey domains) (3;58).

15

RoB item	Low RoB	High RoB
Sequence generation	It will be considered adequate if a random approach in the sequence generation process referred to a random number table, a random computer-generated number, coin tossing, drawing of lots, shuffling of cards, or throwing of dice. Multicentre trials described as randomised will be considered to have adequate sequence generation.	Date of birth, date of inclusion or admission, or record number of clinic/hospital is considered inadequate.
Allocation concealment	It will be considered adequate if there were no reasons to expect that the investigators responsible for inclusion were able to suspect which treatment was next. Both sequentially numbered, sealed, opaque envelopes and a central randomisation are considered adequate.	It will be regarded as inadequate if there is reason to expect that the investigators were able to suspect which treatment was next.
Blinding of patients, personnel, and outcome assessors	It will be considered adequate if the trials describe double-blinding.	It will be considered as inadequate if no blinding is described.
Incomplete outcome data	It will be considered adequate if missing data are ~~few and~~ distributed equally between intervention and control group. Further outcome data will be deemed adequate if data have been imputed using an appropriate technique and analyses based on the ITT ~~analysis~~ population. ~~was applied~~.	It will be considered inadequate if it is unclear how many patients are included in final analyses. Further, it is considered inadequate if no imputation technique is applied or if it is unclear how extensive the missing dataset is (i.e., unclear how many patients withdrew).
Selective reporting	It will be considered adequate if the chosen efficacy outcome (ACR20, ACR50 and/or ACR70) is reported in accordance with the usual contemporary RA protocols and reported at all specified time-points if more than one time point exists.	It will be considered inadequate if the chosen efficacy outcome (ACR20, ACR50 and/or ACR70) is not reported in accordance with the usual contemporary RA protocols, or is not reported at all specified time-points if more than one time point exists.

Table 3: *The Cochrane Risk of Bias tool.*

Other risk of bias components

Funding will be registered according to funding source, as described in **Figure 3** (12). Funding includes provision of manpower (authorship, statistical analysis, or other assistance), study materials (drug, placebo, assay kits, or similar materials), or grants (13). For-profit organisations will be defined as companies that might acquire financial gain or loss depending on the outcome of the trial. Trials partly financed by for-profit agencies will be registered as co-financed.

As an extension to the funding aspect, we will assess whether conflict of interest is reported as "none," denoted by "Yes" or "No."

Figure 3: Funding sources.

To further assess methodological quality, we will note ("Yes" or "No") whether a flow-chart is publicly available.

Single or multicentre status will be determined through review (9). A trial will be considered a multicentre trial if more than one centre is involved. In case of missing information, trials will be classified as multicentre when there is reporting of both several ethics committees and different affiliations of authors. On the other hand, if the report stated only a single ethics committee and a single author affiliation, the trial will be classified as a single centre, unless other information calls for multicentre.

The geographical trial setting will be noted based on the continents participating in the trial: North America, Europe, South America, Asia, Africa and Australia. All the continents involved in the study will be registered.

RoB item	Low RoB	High RoB
Funding	No funding and not-for-profit funding will be considered as low RoB.	For-profit funding and co-financed funding will be considered high RoB. If funding is not reported, it will also be considered high RoB.
Conflict of interest	If conflict of interest is reported as "none," it will be considered low RoB.	If conflict of interest is not reported as "none," it will be considered high RoB.
Flowchart	If a flowchart is publicly available, it will be considered low RoB.	If a flowchart is not publicly available, it will be considered high RoB.
Number of study locations	It will be considered low RoB if more than one centre participated in the trial.	It will be considered high RoB if only one centre participated in the trial, or if it is unclear how many centres participated.

Table 4: *Risk of bias (RoB) components currently not included in the Cochrane RoB tool.*

Two reviewers will independently evaluate eligibility, risk of bias, assessment of ITT/mITT, and handling of missing data. Disagreements will be resolved by discussion until consensus.

Data synthesis

Data synthesis will assess ACR20 data at primary endpoint. We will present differences among different strata by estimating the ratio of odds ratios (ROR). We will assume that the observed number of responders in each arm of each trial has a binomial distribution. Accordingly, intervention effects will be modelled as log-odds ratios and outcomes will be coded so that ORs > 1 correspond to beneficial intervention effects. We will estimate the odds ratio (OR) of trials with the given characteristic using random-effects meta-analyses. For each drug corresponding to a meta-analysis, we will derive the difference between pooled estimates from trials with different trial characteristics (e.g., different ITT approaches). Formal tests of interaction between ITT analysis and estimated treatment benefits (active compared to control) will be performed using the following statistical methodology. After identifying a all trials available for the different biological and targeted agents, we will record characteristics of individual studies (ITT: y/n; Type of ITT: ITT/mITT/AO/PP/Other/Unclear; Type of modification: Treatment/Post baseline assessment/Combined/Other; Handling of missing data in the trial: LOCF/BOCF/Other/Unclear; Handling of NRI: All wd/Lack Of Efficacy wd/Other wd/Unclear) and compare treatment effects

within each biological or targeted agent. <u>As characteristics will be assessed on trial level, analyses will assess comparison on trial level and not within trials.</u>

We will fit *empirical Bayesian* hierarchical bias models using the generalized linear mixed models (GLMM) (59). Mean intervention effects may differ among trials with and without the reported study characteristic. Variation in bias among trials within biologics or targeted agents' trials is quantified and adjusted for with a fixed factor in the model. The GLMMs, like linear mixed models, assume normal (Gaussian) random effects. Conditional on these random effects, data can have any distribution in the exponential family. The exponential family comprises many of the elementary discrete and continuous distributions. The analyses will be performed using the GLIMMIX procedure in SAS (v. 9.2; SAS Institute Inc., Cary, NC, USA) (60). The syntax is similar to that of the MIXED procedure and includes CLASS, MODEL, and RANDOM statements. Using the GLIMMIX procedure, we will perform mixed-effects logistic regression with an arm-based, random-effects model within an *empirical Bayes* framework:

```
Proc Glimmix;
Class Trial Drug Group Characteristic;
Model Counts/Total = Group Characteristic Group*Characteristic / Solution;
Random Trial Trial*Group;
Lsmeans Group*Characteristic / cl ilink;
Run;
```

The PROC GLIMMIX statement invokes the procedure. The CLASS statement instructs the procedure to treat the variables Trial, Drug, Group, and Characteristic as classification variables. The MODEL statement specifies the response variable as a sample proportion using the r/N syntax: Counts/Total corresponds to Y_{iA}/N_{iA} for observations from Group A and to Y_{iB}/N_{iB} or observations from Group B. The SOLUTION option in the MODEL statement requests a listing of the fixed-effects parameter estimates. Because of the response/N syntax, the GLIMMIX procedure defaults to the binomial distribution, with the default logit link. The RANDOM statement specifies that the linear predictor contains intercept terms that randomly vary at the level of the Trial and Trial×Group effects. The default estimation technique in GLMMs is residual pseudo-likelihood (RSPL) with a

subject-specific expansion. The default optimization technique for GLMMs is the Quasi-Newton method. Because a residual likelihood technique is used to compute the objective function, only the covariance parameters are participating in the optimization.

The LsMeans statement requests the least-squares means of the interaction between group (active vs. control) and the individual study characteristic effect on the logit scale. The CL option requests their confidence limits. The ILINK option adds estimates, standard errors, and confidence limits on the mean (probability) scale.

For secondary analyses we will consider possible interaction and confounding by other bias items presented in **Table 3 and 4**.

DISCUSSION

Biased results from RCTs ultimately put the patients at risk for being treated with pharmaceuticals with questionable efficacy and which may cause harm. Taking into account the expenses of accompanying RA treatment, this study is not only biomedical but also a socioeconomic necessity.

The term mITT is used to describe different methods for excluding participants post-randomised from analysis, thereby affecting and disregarding not only the ITT principle but also—and more importantly—the overriding purpose of ITT. Post-randomisation exclusions are known to induce bias, and theoretically mITT will introduce bias (23;24). Our study aim to establish if the bias is of practical concern, and focuses on the direction and magnitude of bias associated with mITT analyses. This study will present arguments as to why mITT approximates ITT or point to the problems concerning the use of mITT. As the term mITT embraces a broad notion of trials, we will delve into how the different types of modification influence effect size. This study may come out with neutral findings—which would not imply that overall bias associated with mITT analyses can be excluded, but may indicate that our study lacks the statistical power necessary to detect the bias. If some form of mITT can substitute ITT, guidelines regarding the use of mITT should be issued. In general this study examines many determinants, and therefore a risk of type I errors due

to multiple comparisons exists and results must be interpreted carefully regardless of statistical significance (61).

This study is limited by the lack of agreement in how ITT and mITT are defined. Our mITT definition and categorisation is based on deviations described in the literature but have some shortcomings; e.g. in cases where only one post-baseline visit is required the mITT category post-baseline assessment will correspond to a completer's analysis.

As in other meta-epidemiological studies, we are limited by the many sources of heterogeneity, e.g. differences in disease duration, type of RA population and intervention dose. As meta-epidemiological studies concerns methodology and do not aim at establishing the empirical evidence for an intervention effect, this underlying premise of heterogeneity can be viewed as acceptable. However, heterogeneity should always be borne in mind when interpreting results.

Our primary objective is to examine whether mITT is associated with different effect sizes, implying empirical evidence for bias in treatment effects. ITT prevents attrition bias when evaluating treatment assignment but may not provide a true estimate of treatment effect if some patients are non-adherent (62). As the term "bias" comprises deviation from the true intervention effect, it can be perceived as misleading to regard systematic errors in treatment effect between mITT and ITT analyses as "bias," given that ITT analysis may fail to provide a true evaluation of the intervention effect. However, ITT analysis is recommended as the least biased way to estimate intervention effects (7) and concerns regarding the systematic errors between mITT and ITT remain, regardless of terminology. This project builds on the premise that the trials included are otherwise less prone to bias, although there is no guarantee that recent trials on biologics and targeted interventions will be at low risk of bias.

This study may point to potential bias and disadvantages in the handling of missing data in RCTs, otherwise known for having a low risk of bias compared with other study designs (63). SI has been criticised on a theoretical level, but its implication on efficacy outcomes in RA trials is uncharted. Accordingly, this study may provide empirical evidence that can support or contradict existing

critics. Regardless of our findings one should always be careful when interpreting results from trials where data are missing and consider the reasons for missing data and potential impact on effect estimates(7;64).

The study examines potential bias associated with industry funding. It may prove difficult to assess bias, as most RA trials concerning biological and targeted agents have some degree of industry input. Being unable to reject industry bias and unable to estimate the influence, direction, and magnitude of such, the validity of trial results in this industry-permeated field of research is open to conjecture (14).

Dissemination

First author Anna Dossing will draft a paper describing the systematic review; the meta-epidemiological study will be disseminated by peer-review publication and conference presentations.

HISTORY

Protocol first published: 11. December 2013

CONTRIBUTION OF AUTHORS

All authors fulfil the ICMJE guidelines for authorship. All authorsAD, ST, DEF, CG, JB, BBH, HB, RC participated in the conception and design of this protocol. RC provided statistical advice for the design and analysis. AD drafted the manuscript. All authorsAD, ST, DEF, CG, JB, BBH, HB, RC critically reviewed the manuscript for important intellectual content and approved the final version.

FUNDING

This research received grants from the Michaelsen foundation. No sponsor was involved in study design, and no sponsor will have authority in the collection, management, analysis, and interpretation of data. Writing of the report and the decision to submit the results for publication

DECLARATION OF INTEREST

The authors declare no conflicts of interest.

ABBREVIATIONS

ACR	American College of Rheumatology
AO	As observed
bDMARDs	Biological disease-modifying anti-rheumatic drugs
BOCF	Baseline Observation Carried Forward
CRP	C-reactive protein
csDMARDs	Conventional synthetic disease-modifying anti-rheumatic drugs
DAS28	European Disease Activity Score
DMARDs	Disease-modifying anti-rheumatic drugs
EMA	European Medicines Agency
ESR	Erythrocyt sedimentation rate
EULAR	European League Against Rheumatism
FDA	US Food and Drug Administration
GCP	Good clinical practice
GLMM	Generalized linear mixed models
ICTRP	WHO Clinical trials Portal
IR	Inadequate response
ITT	Intention-To-Treat
LOCF	Last Observation Carried Forward
MAR	Missing at random
MCAR	Missing completely at random
MI	Multiple imputation

mITT	modified Intention-To-Treat
MNAR	Missing not at random
MTX	Methotrexat
NRI	Non-responder imputation
OR	Odds ratio
PP	Per protocol
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RA	Rheumatoid Arthritis
RCT	Randomised controlled trial
RF	Rheumatoid factor
RoB	Risk of bias
ROR	Ratio of odds ratios
RSPL	Residual pseudo-likelihood
sDMARDs	Synthetic disease-modifying anti-rheumatic drugs
SI	Single imputation
tsDMARD	Targeted synthetic disease-modifying anti-rheumatic drugs
wd	Withdrawal patients

FIGURE LEGENDS

Figure 1: Overview of modified intention-to-treat (mITT) categories.

The four categories are based on the most common deviations described in the literature.

Figure 2: Data extraction flowchart.

csDMARD: Conventional synthetic disease-modifying anti-rheumatic drugs. bDMARD: Biological disease-modifying anti-rheumatic drugs. tsDMARD: Targeted synthetic disease-modifying anti-rheumatic drugs. IR: Inadequate response. ACR20, 50, 70: American College of Rheumatology 20%, 50%, 70% improvement in disease activity respectively. DAS28: European Disease Activity Score. ITT: Intention-To-Treat. mITT: modified Intention-To-Treat. AO: As observed. PP: Per protocol.

LOCF: Last Observation Carried Forward. BOCF: Baseline Observation Carried Forward. NRI: Non-responder imputation. Wd: Withdrawal patients

Figure 3: Funding sources.

<u>REFERENCES</u>

(1) Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ 2001 Jul 7;323(7303):42-6.

(2) Sterne JA, Juni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. Stat Med 2002 Jun 15;21(11):1513-24.

(3) Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008 Mar 15;336(7644):601-5.

(4) Savovic J, Jones H, Altman D, Harris R, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. Health Technol Assess 2012 Sep;16(35):1-82.

(5) Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

(6) Gluud LL. Bias in clinical intervention research. Am J Epidemiol 2006 Mar 15;163(6):493-501.

(7) Higgins J, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 ed. The Cochrane Collaboration; 2011.

(8) Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med 2001 Dec 4;135(11):982-9.

(9) Dechartres A, Boutron I, Trinquart L, Charles P, Ravaud P. Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study. Ann Intern Med 2011 Jul 5;155(1):39-51.

(10) Bafeta A, Dechartres A, Trinquart L, Yavchitz A, Boutron I, Ravaud P. Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. BMJ 2012;344:e813.

(11) Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ 2013;346:f2304.

(12) Montedori A, Bonacini MI, Casazza G, Luchetta ML, Duca P, Cozzolino F, et al. Modified versus standard intention-to-treat reporting: are there differences in methodological

quality, sponsorship, and findings in randomized trials? A cross-sectional study. Trials 2011;12:58.

(13) Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 2003 Aug 20;290(7):921-8.

(14) Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev 2012;12:MR000033.

(15) Bero LA. Why the Cochrane risk of bias tool should include funding source as a standard item. Cochrane Database Syst Rev 2013;12:ED000075.

(16) Sterne JA. Why the Cochrane risk of bias tool should not include funding source as a standard item. Cochrane Database Syst Rev 2013;12:ED000076.

(17) Shrier I, Steele RJ, Verhagen E, Herbert R, Riddell CA, Kaufman JS. Beyond intention to treat: what is the right question? Clin Trials 2014 Feb;11(1):28-37.

(18) Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say? Clin Trials 2007;4(4):350-6.

(19) Kruse RL, Alper BS, Reust C, Stevermer JJ, Shannon S, Williams RH. Intention-to-treat analysis: who is in? Who is out? J Fam Pract 2002 Nov;51(11):969-71.

(20) Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999 Sep 11;319(7211):670-4.

(21) Baron G, Boutron I, Giraudeau B, Ravaud P. Violation of the intent-to-treat principle and rate of missing data in superiority trials assessing structural outcomes in rheumatic diseases. Arthritis Rheum 2005 Jun;52(6):1858-65.

(22) Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: systematic review. BMJ 2010;340:c2697.

(23) Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. Int J Epidemiol 2005 Feb;34(1):79-87.

(24) Nuesch E, Trelle S, Reichenbach S, Rutjes AW, Burgi E, Scherer M, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. BMJ 2009;339:b3244.

(25) Altman DG, Bland JM. Missing data. BMJ 2007 Feb 24;334(7590):424.

(26) Boers M. Missing data in trials: do we have to keep carrying the last observation forward? Arthritis Rheum 2008 Jan 15;59(1):2-3.

(27) EMA. Guideline on missing data in confirmatory clinical trials. London: European Medicines Agency (EMA); 2010.

(28) White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ 2011;342:d40.

(29) Shao J, Jordan DC, Pritchett YL. Baseline observation carry forward: reasoning, properties, and practical issues. J Biopharm Stat 2009 Jul;19(4):672-84.

(30) Molnar FJ, Hutton B, Fergusson D. Does analysis using "last observation carried forward" introduce bias in dementia research? CMAJ 2008 Oct 7;179(8):751-3.

(31) Altman DG. Missing outcomes in randomized trials: addressing the dilemma. Open Med 2009;3(2):e51-e53.

(32) Baron G, Boutron I, Giraudeau B, Ravaud P. Reporting of radiographic methods in randomised controlled trials assessing structural outcomes in rheumatoid arthritis. Ann Rheum Dis 2007 May;66(5):651-7.

(33) Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.

(34) FDA. Guidance for Industry, Clinical Development Programs for Drugs, Devices, and Biological Products for the Treatment of Rheumatoid Arthritis (RA). 1999.

(35) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988 Mar;31(3):315-24.

(36) Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995 Jan;38(1):44-8.

(37) Nielen MM, van SD, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004 Feb;50(2):380-6.

(38) Dahlqvist SR. Rheumatoid arthritis increased the risk for myocardial infarction in women. ACP J Club 2003 Sep;139(2):50.

(39) Scott DL, Symmons DP, Coulton BL, Popert AJ. Long-term outcome of treating rheumatoid arthritis: results after 20 years. Lancet 1987 May 16;1(8542):1108-11.

(40) Welsing PM, van Gestel AM, Swinkels HL, Kiemeney LA, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum 2001 Sep;44(9):2009-17.

(41) van DH, van AJ, Lard LR, Visser K, Ronday HK, Hulsmans HM, et al. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. Arthritis Rheum 2007 May;56(5):1424-32.

(42) Smolen JS, Landewe R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. Ann Rheum Dis 2010 Jun;69(6):964-75.

(43) Furst DE, Keystone EC, Braun J, Breedveld FC, Burmester GR, De BF, et al. Updated consensus statement on biological agents for the treatment of rheumatic diseases, 2011. Ann Rheum Dis 2012 Apr;71 Suppl 2:i2-45.

(44) Tugwell P, Singh JA, Wells GA. Biologicals for rheumatoid arthritis. BMJ 2011;343:d4027.

(45) Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis 2013 Oct 25.

(46) Strand V, Sokolove J. Randomized controlled trial design in rheumatoid arthritis: the past decade. Arthritis Res Ther 2009;11(1):205.

(47) Stein CM, Pincus T. Placebo-controlled studies in rheumatoid arthritis: ethical issues. Lancet 1999 Jan 30;353(9150):400-3.

(48) EMA. Guideline on clinical investigation of medicinal products other than NSAIDs for treatment of rheumatoid arthritis. 2011.

(49) FDA. Guidance for Indusry, Rheumatoid Arthritis: Developing Drug Products for Treatment. 2013.

(50) Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995 Jun;38(6):727-35.

(51) EMA. Points to consider on clinical investigation of medicinal products other than NSAIDS for treatment of rheumatoid arthritis. London: European Medicines Agency (EMA); 2003.

(52) EMA. Draft guideline on clinical investigation of medicinal products other than non-steroidal anti-inflammatory drugs for treatment of rheumatoid arthritis. London: European Medicines Agency (EMA); 2011.

(53) Aletaha D, Landewe R, Karonitsch T, Bathon J, Boers M, Bombardier C, et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Arthritis Rheum 2008 Oct 15;59(10):1371-7.

(54) Dossing A, Tarp S, Furst DE, Gluud C, Wells GA, Beyene J, et al. Attrition bias in rheumatoid arthritis randomised trials with different modified intention-to-treat approaches: a meta-epidemiological study. PROSPERO 2013 Dec 11;CRD42013006702.

(55) Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med 2009 Aug 18;151(4):W65-W94.

(56) Wells G, Becker JC, Teng J, Dougados M, Schiff M, Smolen J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009 Jun;68(6):954-60.

(57) Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg 2012;10(1):28-55.

(58) Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med 2012 Sep 18;157(6):429-38.

(59) Platt RW, Leroux BG, Breslow N. Generalized linear mixed models for meta-analysis. Stat Med 1999 Mar 30;18(6):643-54.

(60) Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. Cochrane Database Syst Rev 2009;(4):CD007848.

(61) Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004 Jun 15;23(11):1663-82.

(62) Shrier I, Steele RJ, Verhagen E, Herbert R, Riddell CA, Kaufman JS. Beyond intention to treat: What is the right question? Clin Trials 2013 Oct 3.

(63) Jakobsen J C & Gluud C. The Necessity of Randomized Clinical Trials. British Journal og Medicine and Clinical Research 2013 May 1;3(4):1453-68.

(64) Akl EA, Briel M, You JJ, Sun X, Johnston BC, Busse JW, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. BMJ 2012;344:e2809.

Treatment		All patients who have received minimum 1 dose of study drug/placebo.
Post-baseline assessment		All patients who have had at least 1 post-baseline assessment.
mITT	Combined	All patients who had received minimum 1 dose of study drug and had at least 1 post-baseline assessment.
	Other	Exclusion of patients who fit none of the categories above.
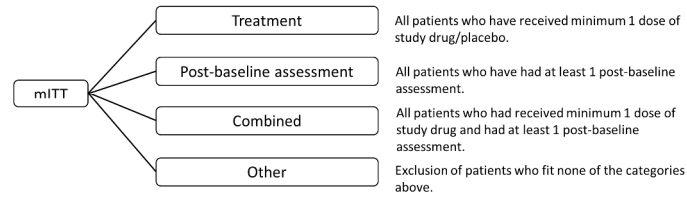
Overview of modified intention-to-treat (mITT) categories. The four categories are based on the most common deviations described in the literature.

254x190mm (300 x 300 DPI)

1. Trial background information
- Author
- Year and journal of publication
- Company study name and number
- Registration number
- Primary outcome

2. Classification of population

Classification of prior treatment and prior treatment responses: csDMARD naive, csDMARD inadequate response (IR), bDMARD or tsDMARD IR.

3. Classification of intervention
- Intervention drug
- Dose
- Form of administration
- Frequency of administration
- Concomitant csDMARD treatment

4. Outcome assessment at primary endpoint

Efficacy: ACR20, 50, and 70 will be assessed. DAS28 data will be extracted as well.
Safety: Overall withdrawal, withdrawal due to adverse events, and number of serious adverse events will be assessed.
We will assess the total number of events and the total number of patients included for analysis.

5. Reported assessment of efficacy analyses

Reported population for efficacy analysis

ITT   mITT   AO   PP   Other   Unclear

See table 2 for elaboration.

6. Reported method of analysis applied?

Yes/No

7. Applied assessment of efficacy analyses

Applied population for efficacy analysis

ITT   mITT   AO   PP   Other   Unclear

See table 2 for elaboration.

Objective 1

8a. Handling of missing data points

Imputation of single missing data points

LOCF   BOCF   Other   Unclear

9a. Handling of withdrawal patients (wd)

When was NRI applied?

All wd   Lack Of Efficacy wd   Other wd   Unclear

Objective 2a

8b. Type of modification applied

mITT

Treatment   Post-baseline assessment   Combined   Other
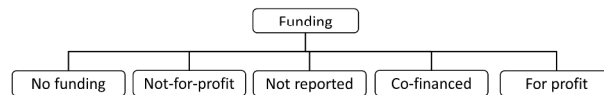
See figure 1 for elaboration.

Objective 2b

9b. No. of modifications applied

Numeric value

Data extraction flowchart. csDMARD: Conventional synthetic disease-modifying anti-rheumatic drugs. bDMARD: Biological disease-modifying anti-rheumatic drugs. tsDMARD: Targeted synthetic disease-modifying anti-rheumatic drugs. IR: Inadequate response. ACR20, 50, 70: American College of Rheumatology 20%, 50%, 70% improvement in disease activity respectively. DAS28: European Disease Activity Score. ITT: Intention-To-Treat. mITT: modified Intention-To-Treat. AO: As observed. PP: Per protocol. LOCF: Last Observation Carried Forward. BOCF: Baseline Observation Carried Forward. NRI: Non-responder imputation. Wd: Withdrawal patients

225x350mm (300 x 300 DPI)

```
                          ┌──────────┐
                          │ Funding  │
                          └──────────┘
     ┌───────────┬────────────┬───────────┬────────────┐
┌──────────┐┌──────────────┐┌────────────┐┌───────────┐┌───────────┐
│No funding││Not-for-profit││Not reported││Co-financed││For profit │
└──────────┘└──────────────┘└────────────┘└───────────┘└───────────┘
```

Funding sources.
254x190mm (300 x 300 DPI)

APPENDIX 1

The Cochrane Library search strategy:

ID	Search
#1	MeSH descriptor: [Recombinant Fusion Proteins] explode all trees
#2	MeSH descriptor: [Antibodies, Monoclonal] explode all trees
#3	MeSH descriptor: [Receptors, Tumor Necrosis Factor] explode all trees
#4	MeSH descriptor: [Receptors, Interleukin-1] explode all trees
#5	MeSH descriptor: [Receptors, Interleukin-6] explode all trees
#6	MeSH descriptor: [Monokines] explode all trees
#7	monoclonal antibody ca2
#8	TNFR-Fc fusion protein
#9	MeSH descriptor: [Interleukin 1 Receptor Antagonist Protein] explode all trees
#10	etanercept
#11	enbrel
#12	infliximab
#13	remicade
#14	adalimumab
#15	humira
#16	D2E7
#17	anakinra
#18	kineret
#19	antril
#20	abatacept
#21	CTLA4Ig
#22	orencia
#23	rituximab
#24	rituxan
#25	idec c2b8
#26	golimumab
#27	simponi
#28	cnto-148
#29	tocilizumab
#30	atlizumab
#31	actemra
#32	roactemra
#33	certolizumab
#34	CDP870
#35	cimzia
#36	"TNFR:Fc":ti,ab,kw (Word variations have been searched)
#37	tofacitinib:ti,ab,kw (Word variations have been searched)

#38	MeSH descriptor: [Janus Kinases] explode all trees
#39	Xeljanz:ti,ab,kw (Word variations have been searched)
#40	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39
#41	MeSH descriptor: [Arthritis, Rheumatoid] explode all trees
#42	Rheumatoid:ti or Rheumatoid:ab (Word variations have been searched)
#43	arthriti*:ti or arthriti*:ab (Word variations have been searched)
#44	#42 and #43
#45	#41 or #44
#46	#40 and #45 in Trials

EMBASE search strategy:

1	abatacept.mp.
2	adalimumab.mp.
3	certolizumab.mp.
4	etanercept.mp.
5	CDP870.mp.
6	golimumab.mp.
7	infliximab.mp.
8	rituximab.mp.
9	tocilizumab.mp.
10	humira.mp.
11	trudexa.mp.
12	orencia.mp.
13	cimzia.mp.
14	enbrel.mp.
15	simponi.mp.
16	rituxan.mp.
17	mabthera.mp.
18	actemra.mp.
19	RoActemra.mp.
20	monoclonal antibodies.mp. or exp Antibodies, Monoclonal/
21	exp Monokines/

22	exp Receptors, Interleukin-1/
23	exp Receptors, Interleukin-6/
24	exp Polyethylene Glycols/
25	exp Immunoglobulin G/
26	exp Immunoconjugates/
27	immunoglobulin fab fragments.mp. or exp Immunoglobulin Fab Fragments/
28	t-lymphocytes.mp. or exp T-Lymphocytes/
29	exp tumor necrosis factor inhibitor/
30	exp interleukin 1 receptor blocking agent/
31	D2E7.mp.
32	anakinra.mp.
33	kineret.mp.
34	antril.mp.
35	CTLA4Ig.mp.
36	idec c2b8.mp.
37	cnto-148.mp.
38	atlizumab.mp.
39	tofacitinib.mp.
40	exp Janus kinase inhibitor/
41	*tumor necrosis factor receptor/dt [Drug Therapy]
42	or/1-41
43	exp Random Allocation/
44	exp Single-Blind Method/
45	exp Double-Blind Method/
46	Placebo.mp.
47	Randomi?ed controlled trial$.mp.
48	rct.mp.
49	Random allocation.mp.
50	Randomly allocated.mp.
51	Allocated randomly.mp.
52	(allocated adj2 random).mp.
53	Single blind$.mp.
54	Double blind$.mp.
55	((treble or triple) adj blind$).mp.
56	Placebo$.mp.

57	or/43-56
58	rheumatoid.ti,ab.
59	*rheumatoid arthritis/
60	58 or 59
61	42 and 57 and 60
62	limit 61 to (book or book series or conference abstract or conference paper or conference proceeding or "conference review")
63	61 not 62

LILACS search strategy:

(tw:(tw:((tw:(rheumatoid)) AND (tw:(etanercept)) OR (tw:(enbrel)) OR (tw:(infliximab)) OR (tw:(remicade)) OR (tw:(adalimumab)) OR (tw:(humira)) OR (tw:(d2e7)) OR (tw:(anakinra)) OR (tw:(kineret)) OR (tw:(antril)) OR (tw:(abatacept)) OR (tw:(ctla4ig)) OR (tw:(orencia)) OR (tw:(rituximab)) OR (tw:(rituxan)) OR (tw:(idec c2b8)) OR (tw:(golimumab)) OR (tw:(simponi)) OR (tw:(cnto-148)) OR (tw:(tocilizumab)) OR (tw:(atlizumab)) OR (tw:(actemra)) OR (tw:(roactemra)) OR (tw:(certolizumab)) OR (tw:(cdp870)) OR (tw:(cimzia)) OR (tw:(tnfr:fc)) OR (tw:(tofacitinib)) OR (tw:(janus kinases)) OR (tw:(xeljanz))) AND db:("LILACS"))) AND (tw:(random*))

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	8
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	8-9
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	9+Appendix 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	10-13
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	10-13
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	13-15
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	16
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	16-17

PRISMA 2009 Checklist

Page 1 of 2

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	13-15
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	16-17
RESULTS			Study protocol, no results
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	18
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	18-19
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Study protocol, no results
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	20

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.* PLoS Med 6(6): e1000097.

PRISMA 2009 Checklist

For more information, visit: **www.prisma-statement.org**.

Page 2 of 2