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 Christensen1

ABSTRACT

Introduction: When participants drop out of randomised clinical trials, as frequently happens, the intention-to-treat (ITT) 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 intention-to-treat (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 (ie, good internal validity) and assessing biological or targeted agents in patients with rheumatoid arthritis, 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 healthcare researchers, policymakers and clinicians. Results of the study will be disseminated by peer-review publication.

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

INTRODUCTION

Meta-epidemiological research

Inadequate quality of trials may distort the results from meta-analyses and systematic reviews. Consequently, meta-epidemiological studies are carried out to quantitatively evaluate bias across many randomised clinical trials (RCTs) in different meta-analyses. 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.
The ITT 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 postrandomisation compared to trials reporting on the ITT population.12 Although postrandomisation exclusion is known to produce bias,23 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 data sets due to withdrawal.20 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.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.27 29 LOCF and BOCF have been widely criticised.30–32 and the use of SI risk introducing bias, for example, by minimising variation.27 32 A newer and more promising imputation method is multiple imputation (MI), which involves creating several different plausible imputed data sets 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 failure27 34 (also referred to as non-responder imputation (NRI)). NRI is applied inconsistently to different withdrawal populations, for example, 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 populations27 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.25 26 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 carried out. 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 RA
RA is a chronic inflammatory autoimmune disease characterised by joint swelling and joint tenderness with...
Disease Activity Score (DAS28). The DAS28 score response criteria and/or variations of the European using the American College of Rheumatology (ACR) may contribute to attrition bias. Rescue therapy challenges the ITT principle and encountered as withdrawal patients in the trials protocols; patients who receive rescue therapy are commonly associated with arthritis. 

Treatment of RA encompasses multiple interventions. Though RA has been shown to progress over time despite treatment, early therapeutic intervention improves clinical outcomes and reduces the accrual of joint damage and disability. RA can be managed using disease-modifying antirheumatic drugs (DMARDs). DMARDs form two major classes: synthetic chemical compounds (sDMARDs) and biological agents (bDMARDs). sDMARDs comprise the conventional DMARDs (csDMARDs) (eg, methotrexat (MTX) and sulfasalazin) and the new targeted sDMARDs (tsDMARDs); for example, 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. The DMARDs of interest are summarised in Table 1.

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 patients with RA. Rescue therapy (eg, 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. Since 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). The DAS28 score discriminates between high and low disease activity. 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 following: (1) patient assessment; (2) physician assessment; (3) erythrocyte sedimentation rate; (4) pain scale; and (5) functional questionnaire: 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.

**Objective**

Our primary objective is to examine

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

2. 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 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 (CRD42013006709), our manuscript conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for reporting systematic reviews and meta-analyses.

**Eligibility criteria**

All RA RCTs concerning EMA-approved 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 end point.
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 online supplementary appendix 1 for the EMBASE, the Cochrane Library and LILACS search strategies. The WHO Clinical trials Portal (ICTRP), clinical trials.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 “cimzia”[All Fields] OR “cimzia”[All Fields] OR “cimzia”[All Fields])) OR (“tocilizumab”[Supplementary Concept] OR “tocilizumab”[All Fields] OR “tocilizumab 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 arthritis[“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 (ie, safety, efficacy or both) and the name of the primary outcome (eg, 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 erythrocyte sedimentation rate (ESR), we will extract only DAS28 (ESR), as both outcomes are in agreement. For safety outcomes, we will assess overall withdrawals, withdrawals due to adverse events and number of serious adverse events.

ITT 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 (ie, analyses conducted on all patients who completed the study) and patients with imputed outcome data (ie, imputed data.
from all withdrawal patients). The specifics of each assessment are presented in table 2.22

The CONSORT statement has proposed equivalence between PP analysis and mITT covering exclusions of participants who did not adequately adhere to the protocol57; 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: (1) patients who withdrew due to lack of efficacy; (2) all patients who withdrew or (3) another defined patient group that withdrew but fit neither 1 or 2.

When mITT is applied, we will assess the method of modification based on prior described categories (see figure 1).9 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 postbaseline 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 postbaseline assessment. Modifications that fail to fit into the three aforementioned categories will be registered as ‘other.’ The number of modifications applied will be extracted as a numeric value.

A complete data extraction flow chart is presented in figure 2. All data extraction will be carried out on trial level, except classification of intervention (step 3) and outcome assessment at primary end point (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.

### Table 2  Analysis population for efficacy analysis and adverse outcome analysis12 19 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 prespecified in the protocol are included in efficacy analysis and adverse outcome analysis |
| Other               | None of the aforementioned categories fits the analysis population |
| Unclear             | It is unclear which analysis is applied |

ITT, intention to treat; mITT, modified intention to treat.

---

Evaluating methodological quality

The Cochrane RoB tool

The RoB within each trial will be assessed using the RoB tool as recommended by The Cochrane Collaboration (see below).3 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 RoB 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, categorised as low risk (low for all Cochrane components) or high risk (high or unclear for ≥1 Cochrane components).58

Other RoB 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 cofinanced.

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

**Figure 2** Data extraction flow chart. ACR20, 50, 70, American College of Rheumatology 20%, 50%, 70% improvement in disease activity respectively; AO, as observed; bDMARD, biological disease-modifying antirheumatic drugs; BOCF, Baseline Observation Carried Forward; csDMARD, conventional synthetic disease-modifying antirheumatic drugs; DAS28, European Disease Activity Score; IR, inadequate response. tsDMARD, targeted synthetic disease-modifying antirheumatic drugs; ITT, intention-to-treat; LOCF, last observation carried forward; mITT, modified intention-to-treat; NRI, non-responder imputation; PP, per protocol; Wd, withdrawal patients.
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. 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.

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.

### Table 3  The Cochrane risk of bias tool

| 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 regarded as 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 data set is (ie, unclear how many patients withdrew). 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. |
| 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. |                                                                                                                                                                                                                             |

ACR20, ACR50 and/or ACR70, American College of Rheumatology 20%, 50%, 70% improvement in disease activity respectively; ITT, intention to treat; RA, rheumatoid arthritis; Rob, Risk of bias.

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

| RoB item                  | Low RoB                                                                                           | High RoB                                                                                                                                                                                                                     |
|---------------------------|--------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Funding                   | No funding and not-for-profit funding will be considered as low RoB. For-profit funding and cofinanced 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. | If conflict of interest is not reported as ‘none,’ it will be considered high RoB. If conflict of interest is reported as ‘none,’ it will be considered low RoB.                                                                 |
| Flow chart                | If a flow chart is publicly available, it will be considered low RoB. If a flow chart is not publicly available, it will be considered high RoB. | If a flow chart 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.                                                                                             |
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 $Y_{IA}/N_A$ for observations from Group A and to $Y_{IB}/N_B$ or observations from Group B. The SOLUTION option in the MODEL statement requests a listing of the fixed-effects parameter estimates. Owing to 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 pseudolikelihood with a subject-specific expansion. The default optimisation technique for GLMMs is the Quasi-Newton method. Since a residual likelihood technique is used to compute the objective function, only the covariance parameters are participating in the optimisation.

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, SEs and confidence limits on the mean (probability) scale.

For secondary analyses we will consider possible interaction and confounding by other bias items presented in tables 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 socio-economic necessity.

The term mITT is used to describe different methods for excluding participants postrandomised from analysis, thereby affecting and disregarding not only the ITT principle but also—and more importantly—the overriding purpose of ITT. Postrandomisation 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 1 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; for example, in cases where only one postbaseline visit is required the mITT category postbaseline assessment will correspond to a completer’s analysis.

As in other meta-epidemiological studies, we are limited by the many sources of heterogeneity, for example, 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 attribution bias when evaluating treatment assignment but may not provide a true estimate of treatment effect if some patients are non-adherent.17 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 effects7 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 RoB compared with other study designs.62 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 63

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

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

Contributors AD, ST, DEF, CG, JB, BBH, HB and 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 and RC critically reviewed the manuscript for important intellectual content and approved the final version.

Funding
This research received grants from the Michaelsen foundation. Musculoskeletal Statistics Unit, The Parker Institute, is supported by grants from the Oak Foundation. The Copenhagen Trial Unit is funded by the Danish state.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES
1. Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. BMJ 2001;323:42–6.
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;21:1513–24.
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;336:601–5.
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 from meta-epidemiological studies. Health Technol Assess 2012;16:1–82.
5. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5029.
6. Gluud LL. Bias in clinical intervention research. Am J Epidemiol 2006;163:493–501.
7. Higgins J, Altman DG, Sterne JAC. Chapter 8: assessing risk of bias in included studies. In: Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 edn. The Cochrane Collaboration, 2011.
8. Kjaergard LL, Villumsen J, Gluud C. Reported methodological quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med 2001;135:982–9.
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;155:39–51.
