

Major Mistakes and Errors in the Use of Trial Sequential Analysis in Systematic Reviews or Meta-analyses – Protocol for a Systematic Review

Christian Gunge Riberholt (✉ christian.riberholt@regionh.dk)

Rigshospitalet <https://orcid.org/0000-0002-6170-1869>

Markus Harboe Olsen

Copenhagen University Hospital Rigshospitalet

Joachim Birch Milan

Copenhagen University Hospital Rigshospitalet

Christian Gluud

Copenhagen University Hospital Rigshospitalet

Protocol

Keywords: Meta analysis, methodology, systematic review, Trial Sequential Analysis

Posted Date: September 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-900530/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Adequately conducted systematic reviews with meta-analyses are considered the highest level of evidence and thus directly defines many clinical guidelines. However, the risk of type I and II errors in meta-analyses are substantial. Trial Sequential Analysis is a method for controlling these risks. Erroneous use of the method might lead to research waste or misleading conclusions.

Methods: The current protocol describes a systematic review aimed to identify common and major mistakes and errors in the use of Trial Sequential Analysis by evaluating published systematic reviews and meta-analyses that include this method. We plan to include all studies using Trial Sequential Analysis published from 2018 to 2021, an estimated 400 to 600 publications. We will search Medical Literature Analysis and Retrieval System Online (MEDLINE) and the Cochrane Database of Systematic Reviews (CDSR), including studies with all types of participants, interventions, and outcomes. The search will begin in July 2021. Two independent reviewers will screen titles and abstracts, include relevant full text articles, extract data from the studies into a predefined checklist, and evaluate the methodological quality of the study using the AMSTAR 2 (Assessing the methodological quality of systematic reviews).

Discussion: This protocol follows the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P). The identified mistakes and errors will form the basis of a reviewed guideline for the use of Trial Sequential Analysis. Appropriately controlling for type I and II errors might reduce research waste and improve quality and precision of the evidence that clinical guidelines are based upon.

Introduction

Adequately conducted systematic reviews with meta-analyses are considered the highest level of evidence within evidence-based medicine [1, 2]. The number of systematic reviews and meta-analysis guiding clinical practice and decision making has been steadily increasing during the past four decades [3].

Despite their place at the top of the hierarchical research pyramid, conventional meta-analyses are still at risk of type I error (*alpha*) due to results reaching significance by chance and type II errors (*beta*) due to results not reaching significance even when an effect is present. The risk of these errors is generally accepted at consensus-based levels (typically 5% for type I and 10–20% for type II errors), but may increase beyond those levels due to publication bias, poor trial designs, data heterogeneity and poorly conducted or inadequately powered meta-analyses with multiple significance testing [1, 4, 5]. The investigated effect of a meta-analysis can reach significance even though the effect might be so small that it is not clinically relevant [6].

Several tools exist for controlling for non-random causes of type I and II errors as rigorously described by the Cochrane Handbook [2]. However, little emphasis has been put on mitigating the purely random causes of type I and II errors [7]. As an example, correction for multiplicity issues due to use of several outcomes has historically been underprioritised [5, 8]. As shown by Turner, Bird and Higgins in 2013, 70% of the meta-analyses reported in Cochrane Reviews had less than 50% power to detect even a relatively large 30% relative risk reduction [9]. Moreover, there is an increased risk of an exaggerated intervention benefit in small trials due to reporting bias or methodological flaws [10].

Pooling randomised clinical trials in a meta-analysis may introduce heterogeneity due to differences in design, interventional protocols, included participants, and outcomes [11, 12]. This heterogeneity can be an advantage as it may reflect the naturally occurring variation in clinical practice compared to trials with very strict or similar protocols, thereby mimicking real-world treatment [1]. In a meta-analytic setting, however, heterogeneity needs to be adequately examined and considered by Trial Sequential Analysis [13, 14].

Trial Sequential Analysis was developed to control the risk of type I and II errors in meta-analysis. It can be used to estimate the diversity-adjusted required information size (DARIS or the 'meta-analytic sample size') in random-effects meta-analysis, which serves to indicate whether the chosen power level was reached with the meta-analysis. If adequate power was not reached, DARIS may guide the scaling of future trials [14]. It also establishes when firm evidence is reached for or against a specific intervention.

Further, Trial Sequential Analysis can establish futility boundaries and thus indicate when non-significant results are due to lack of intervention effect and not lack of power [13, 15].

The Trial Sequential Analysis is a sequential method using α -spending monitoring boundaries [16]. For dichotomous outcomes, one needs a proportion of participants with the outcome in the control group (P_c), an a priori chosen anticipated intervention effect (relative risk reduction (RRR)), alpha level, beta level, and the diversity (D^2) of the trials included in the Trial Sequential Analysis for the calculation of the DARIS in a random-effects model meta-analysis [1, 15]. For continuous outcomes, one needs a minimally relevant clinical difference, the associated variance (squared standard deviation), the alpha level, beta level, and the diversity (D^2) of the trials included in the Trial Sequential Analysis for the calculation of the diversity-adjusted required information size (DARIS or the 'meta-analytic sample size') in a random-effects model meta-analysis [17].

Furthermore, Trial Sequential Analysis can advantageously be used alongside the Grading of Recommendations Assessment, Development and Evaluation (GRADE). By calculating the DARIS, the reviewers are informed of the power of their analysis and can therefore grade imprecision in a more informative manner [18]. Also, the analysis can supply the reviewers with a *trial sequential adjusted confidence interval* to evaluate inconsistency alongside the heterogeneity [16].

To date, numerous systematic reviews have used Trial Sequential Analysis since it was first presented at the beginning of this millennium [19]. As with all methods, the Trial Sequential Analysis can be misused and misinterpreted. A rigorous process starting when writing the protocol through to the reporting phase of the results in the review is necessary. Predefined parameters such as alpha level, beta level (and power), relative risk reduction, minimally relevant clinical difference, and heterogeneity can largely affect the results of the analysis and should therefore be enclosed in pre-published or registered protocols prior to searching for literature for the systematic review. Failure to do so might ultimately alter the conclusion of the meta-analysis and thereby directly misguide clinical practice [5, 16, 20].

Objective

In this review, we aim to systematically evaluate the use of Trial Sequential Analysis in existing systematic reviews and meta-analyses. Specifically, we seek to evaluate how the authors prepared and conducted their Trial Sequential Analysis, and interpreted their results in the assessment of imprecision in the obtained meta-analytic results. We want to identify the most common major mistakes and errors in order to update recommendations for a more proper use of the Trial Sequential Analysis Programme in future systematic reviews.

Methods

Criteria for considering studies for this quality assessment study

As this is a methodological review examining the use of Trial Sequential Analysis in systematic reviews or in meta-analyses, there are only a few criteria for considering eligible reviews. This protocol adheres to the reporting guidelines PRISMA-P (**Supplemental material**) [21].

Types of studies

This review will include peer reviewed publications of systematic reviews with meta-analyses or of meta-analysis of randomised clinical trials that have included a Trial Sequential Analysis and analysed at least two randomised clinical trials. A meta-analysis is a statistical approach for combining data, while a systematic review is a detailed, organised, and transparent method of gathering, appraising and synthesising data to answer a well-defined question [2]. Most systematic reviews contain meta-analyses. The included studies must at least include two randomised clinical trials in at least one meta-analysis and Trial Sequential Analysis. Only the most recent studies will be included, i.e. those published from 2019 to 2021. For practical reasons, only articles in English will be included in the study. We expect that we will identify 400 to 600 relevant publications.

Types of participants

We accept all participants of any race, sex, or age with any disease or condition for this review.

Types of interventions

We accept all types of intervention for this review.

Types of outcomes

All outcomes, dichotomous or continuous, are accepted for this review if they are analysed using both meta-analysis and Trial Sequential Analysis.

Search strategy

The following databases will be sought:

- Medical Literature Analysis and Retrieval System Online (MEDLINE)
- The Cochrane Database of Systematic Reviews (CDSR)

Keywords used in the search strategy:

- Trial Sequential Analysis OR TSA
- Systematic Review OR Meta-analysis

The preliminary search strategy can be found in **Supplemental material**.

The references of included reviews will be screened to identify further trials of interest.

Selection of studies

Two authors (CGR and MHO) will independently screen the title and abstract using the web-based application Covidence (www.covidence.org, Melbourne, Australia) [22]. All relevant full-text articles will be retrieved and screened for eligibility, and reasons for exclusion will be recorded. Any discrepancy will be resolved through discussion. If an agreement is not reached, a third author (CG) will resolve the disagreement. Trial selection will be shown through a PRISMA flow chart following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (**Supplemental material**). References from the included studies will be reviewed by an expert in the field (CG) for any missing studies.

Methodological quality of the systematic reviews and meta-analyses

Two authors will independently evaluate the methodological quality in all included systematic reviews and meta-analyses using the AMSTAR 2 (Assessing the methodological quality of systematic reviews) [23]. The assessment of the methodological quality of the studies will be used to evaluate whether improper use of Trial Sequential Analysis is related to other methodological flaws. Each of the 16 items will be rated, and a final rating of the overall confidence in the results of each study will be given on a scale of confidence as high, moderate, low, or critically low [23]. Any discrepancy will be resolved through discussion. If an agreement is not reached, a third author (CG) will resolve the disagreement.

Extraction of data

Two independent authors will extract data from each included study. After extraction, all data will be compared, consensus will be reached, or a third author will be consulted to resolve disagreement.

General information on each study will be extracted (author, year, number of included trials in the analyses, etc.). Specific data regarding the Trial Sequential Analysis will be extracted systematically through a predefined checklist made in REDCap (Research Electronic Data Capture, University of Kansas, United States) hosted at Rigshospitalet [24–26].

A prespecified extraction checklist containing items about the methodology of Trial Sequential Analysis will be prepared prior to the literature search. By using the Trial Sequential Analysis manual[17] and randomly selecting 10 systematic reviews, including

meta-analysis and Trial Sequential Analysis, the most common and important steps in the analysis has been selected by the review group and synthesised into the extraction checklist. The extraction checklist contains four main categories (identification, content of the pre-registered protocol, content of the systematic review, and results), each of which has relevant questions depending on the type of outcome (dichotomous or continuous) used in the review. For more information on the extraction checklist, see **Supplemental material**.

Assessment of Trial Sequential Analysis results for downgrading for imprecision

We will assess the way authors of systematic reviews or meta-analyses have used Trial Sequential Analysis for downgrading for imprecision in GRADE [27, 28] or by other methods. We will compare the downgrading for imprecision in the included studies with the following method: imprecision in GRADE is downgraded by two levels if the accrued number of participants is below 50% of the DARIS, and one level if between 50% and 100% of DARIS. We will not expect downgrading if the cumulative Z-curve crosses the monitoring boundaries for benefit, harm or futility, or if DARIS is reached. We will examine if such methodology has an impact on the conclusions drawn in the systematic review or meta-analysis.

Data analysis

The most common mistakes will be ranked according to their prevalence and comparisons of the AMSTAR 2 score in groups of studies will be made. The consequences of the most common errors and mistakes found in the literature when using Trial Sequential Analysis will be explained by examples and suggestions on how to correct these. Finally, a guideline for future reviewers will be created from the identified mistakes and errors.

Mistakes and errors will be categorised as related to the protocol, methodology, presentation of the results, and the interpretation. Each mistake and error will, furthermore, be classified as major or minor. Major mistakes or errors are those with the potential to cause a wrong conclusion. This classification will be based on a consensus by the investigators.

Statistical considerations

The number of major or minor mistakes and errors per article will be presented as median and interquartile range. The AMSTAR 2 score for articles without any major mistakes will be compared to those with any major mistake will be compared using either Students t-test or Wilcoxon-signed rank test depending on normality.

The mistakes and errors will be presented as frequencies, with a 95% confidence interval calculated using 1-sample proportions test without continuity correction. The AMSTAR 2 score for the manuscripts where the specific mistake and error was present will be presented as a covariate. We will combine these mistakes and errors based on our recommendations, as errors referring to both protocol and presentation of results might be handled by one recommendation. Both the specified and aggregated frequency will be presented in a table related to the recommendation.

Discussion

This review aims to assess the use of Trial Sequential Analysis in the current body of systematic reviews and meta-analyses. Trial Sequential Analysis offers important pieces of information [4, 15, 29] and provides more stringent planning of how to calculate the DARIS and interpret the imprecision in GRADE [16, 18, 20, 29–33] than present recommendations regarding the calculation of optimal information size and assessment of imprecision in GRADE [27]. However, the method is currently not a mandatory part of Cochrane Reviews. Arguments put forward against the use is that authors of systematic reviews and meta-analysis rarely has the power to start or stop a trial being conducted and that conclusions of meta-analyses should not be driven by statistical testing [34, 35].

Since the introduction of Trial Sequential Analysis, an increasing number of authors have used the Trial Sequential Analysis to control the risk of type I and II errors and thus improve the quality of evidence and the recommendations. However, strict systematic approaches to this analysis are important as it can be misused for a '*cherry-picking*' approach. Therefore, this systematic review of the methodology of other selected systematic reviews is important for understanding and conceptualising

the use of Trial Sequential Analysis. As mistakes and errors in the use of the Trial Sequential Analysis are likely to be found, these will be used to establish a guideline for future reviewers.

Trial Sequential Analysis has been accepted as a supplementary analysis in Cochrane Hepato-Biliary Group systematic reviews (<https://hbg.cochrane.org/information-authors>) and as stated in the Cochrane Handbook for Systematic Reviews of Interventions: “...trial sequential analysis may, however, be used in the context of a prospectively planned series of randomized trials” [36].

Stage of the review at the time of the submission

At the time of submitting this protocol, ten randomly selected systematic reviews with meta-analysis and Trial Sequential Analysis were used to test and improve the extraction checklist. Minor changes to the extraction checklist can occur during the first stage of data extraction. A preliminary search was done on the 9th of July, 2021 and the final search will be conducted after the submission of this protocol.

Abbreviations

AMSTAR 2: Assessing the methodological quality of systematic reviews 2; D2: Diversity; DARIS: Diversity-adjusted required information size; GRADE: Grading of Recommendations Assessment, Development and Evaluation; Pc: Proportion of participants with the outcome in the control group; RRR: Relative risk reduction

Declarations

Ethics and approvals

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Competing interests

The authors' declare that they have no competing interests.

Funding

No external funding was received for this study.

Authors' contribution

CGR, MHO, JBM, and CG are responsible for the conception and design of the study. CGR is the guarantor of the protocol and drafted the manuscript, and all authors approved the final version of the manuscript.

Acknowledgements

We would like to thank Information specialist Sarah Louise Klingenberg at The Cochrane Hepato-Biliary Group, Copenhagen for help in developing the search strategy.

References

1. Garattini S, Jakobsen JC, Wetterslev J, Bertelé V, Banzi R, Rath A, et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. *Eur J Intern Med.* 2016;32:13–21. doi:10.1016/j.ejim.2016.03.020.

2. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.1. Cochrane; 2020. www.training.cochrane.org/handbook.
3. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q.* 2016;94:485–514. doi:10.1111/1468-0009.12210.
4. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol.* 2008;61:763–9. doi:10.1016/j.jclinepi.2007.10.007.
5. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol.* 2014;14.
6. Shrier I, Christensen R, Juhl C, Beyene J. Meta-analysis on continuous outcomes in minimal important difference units: an application with appropriate variance calculations. *J Clin Epidemiol.* 2016;80:57–67. doi:10.1016/j.jclinepi.2016.07.012.
7. Roberts I, Ker K, Edwards P, Beecher D, Manno D, Sydenham E. The knowledge system underpinning healthcare is not fit for purpose and must change. *BMJ.* 2015;350. doi:10.1136/bmj.h2463.
8. Bender R, Bunce C, Clarke M, Gates S, Lange S, Pace NL, et al. Attention should be given to multiplicity issues in systematic reviews. *Journal of Clinical Epidemiology.* 2008;61:857–65. doi:10.1016/j.jclinepi.2008.03.004.
9. Turner RM, Bird SM, Higgins JPT. The impact of study size on meta-analyses: Examination of underpowered studies in Cochrane reviews. *PLoS One.* 2013;8.
10. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med.* 2001;135:982–9. doi:10.7326/0003-4819-135-11-200112040-00010.
11. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21:1539–58. doi:10.1002/sim.1186.
12. Barbateskovic M, Koster TM, Eck RJ, Maagaard M, Afshari A, Blokzijl F, et al. A new tool to assess Clinical Diversity In Meta-analyses (CDIM) of interventions. *J Clin Epidemiol.* 2021;135:29–41. doi:10.1016/j.jclinepi.2021.01.023.
13. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Med Res Methodol.* 2009;9:86. doi:10.1186/1471-2288-9-86.
14. Claire R, Gluud C, Berlin I, Coleman T, Leonardi-Bee J. Using Trial Sequential Analysis for estimating the sample sizes of further trials: example using smoking cessation intervention. *BMC Med Res Methodol.* 2020;20. doi:10.1186/s12874-020-01169-7.
15. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol.* 2008;61:64–75.
16. Wetterslev J, Jakobsen JC, Gluud C. Trial Sequential Analysis in systematic reviews with meta-analysis. *BMC Med Res Methodol.* 2017;17:1–18.
17. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for Trial Sequential Analysis (TSA). www.ctu.dk/tsa/files/tsa_manual.pdf.
18. Castellini G, Bruschetti M, Gianola S, Gluud C, Moja L. Assessing imprecision in Cochrane systematic reviews: A comparison of GRADE and Trial Sequential Analysis. *Syst Rev.* 2018;7:110. doi:10.1186/s13643-018-0770-1.
19. Thorlund K, Wetterslev J, Brok J, Gluud C. Trial sequential analyses of six meta-analyses considering heterogeneity and trial weight (Abstract). In: *Cochrane Colloquium.* Melbourne; 2005. <https://abstracts.cochrane.org/2005-melbourne/trial-sequential-analyses-six-meta-analyses-considering-heterogeneity-and-trial>.
20. Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JPA, Thabane L, et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? *Int J Epidemiol.* 2009;38:276–86. doi:10.1093/ije/dyn179.
21. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. 2015. doi:10.1186/2046-4053-4-1.
22. Covidence®. Covidence systematic review software [Computer program]. www.covidence.org.
23. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ.* 2017;358:4008. doi:10.1136/bmj.j4008.

24. REDCap. <https://www.project-redcap.org/>.
25. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42:377–81.
26. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O’Neal L, et al. The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics.* 2019;95:103208.
27. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence–imprecision. *J Clin Epidemiol.* 2011;64:1283–93. doi:10.1016/j.jclinepi.2011.01.012.
28. Castellini G, Bruschetti M, Gianola S, Gluud C, Moja L. Assessing imprecision in Cochrane systematic reviews: A comparison of GRADE and Trial Sequential Analysis. *Syst Rev.* 2018;7:110. doi:10.1186/s13643-018-0770-1.
29. Bangalore S, Wetterslev J, Pranesh S, Sawhney S, Gluud C, Messerli FH. Perioperative β blockers in patients having non-cardiac surgery: a meta-analysis. *Lancet.* 2008;372:1962–76. doi:10.1016/S0140-6736(08)61560-3.
30. Gartlehner G, Nussbaumer-Streit B, Wagner G, Patel S, Swinson-Evans T, Dobrescu A, et al. Increased risks for random errors are common in outcomes graded as high certainty of evidence. *J Clin Epidemiol.* 2019;106:50–9. doi:10.1016/j.jclinepi.2018.10.009.
31. Nielsen N, Friberg H, Gluud C, Herlitz J, Wetterslev J. Hypothermia after cardiac arrest should be further evaluated-A systematic review of randomised trials with meta-analysis and trial sequential analysis. *Int J Cardiol.* 2011;151:333–41. doi:10.1016/j.ijcard.2010.06.008.
32. Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive–Trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *Int J Epidemiol.* 2009;38:287–98. doi:10.1093/ije/dyn188.
33. Korang SK, Juul S, Nielsen EE, Feinberg J, Siddiqui F, Ong G, et al. Vaccines to prevent COVID-19: a protocol for a living systematic review with network meta-analysis including individual patient data (The LIVING VACCINE Project). *Syst Rev.* 2020;9. doi:10.1186/s13643-020-01516-1.
34. Gluud C, Wetterslev J, Higgins J, Simmonds M, et al. Trial Sequential Analysis or sequential meta-analysis. *Cochrane Scientific Committee.* 2017. https://methods.cochrane.org/sites/default/files/public/uploads/2017_1_18_may_scientific_committee_agenda_docs.pdf. Accessed 15 May 2021.
35. Schmid C, Chandler J, Senn S, Sterne J, Kulinskaya E, Posch M, et al. Should Cochrane apply error-adjustment methods when conducting repeated meta-analyses? *Cochrane Sci Comm.* 2018. https://methods.cochrane.org/sites/default/files/public/uploads/tsa_expert_panel_guidance_and_recommendation_final.pdf.
36. Thomas J, Askie L, Berlin J, Elliott J, Ghersi D, Simmonds M, et al. Prospective approaches to accumulating evidence. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., editors. *Cochrane Handbook for Systematic Reviews of Interventions.* 6.1. Cochrane; 2020. www.training.cochrane.org/handbook.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [20210913Supplementalmaterial.pdf](#)