

Comparison of non-inferiority margins reported in protocols and publications showed incomplete and inconsistent reporting

Olaf M. Dekkers^{a*}, Myriam Cevallos^{b,c*}, Jonas Bühler^b, Antoine Poncet^d, Sabine Allemann^b,
Thomas V. Perneger^d, Matthias Egger^b

* Contributed equally

a Department of Clinical Epidemiology, Leiden University Medical Center, The Netherlands

b Institute of Social & Preventive Medicine (ISPM), University of Bern, Switzerland

c Clinical Trials Unit (CTU) Bern, University of Bern, Switzerland

d Division of Clinical Epidemiology, University Hospitals of Geneva, Geneva, Switzerland

Correspondence to:

Prof. Matthias Egger

Institute of Social & Preventive Medicine (ISPM)

Finkenhubelweg 11

CH-3012 Bern

Switzerland

Tel +41 31 631 35 11

Fax +41 31 631 35 20

Email: egger@ispm.unibe.ch

Word count: Main text 3397 words, abstract 255 words, 1 figure, 3 tables, 27 references.

Abstract

Objectives: To compare non-inferiority margins defined in study protocols and records in trial registries with margins reported in subsequent publications.

Study design and setting: Comparison of protocols of non-inferiority trials submitted 2001 to 2005 to ethics committees in Switzerland and The Netherlands with corresponding publications and registry records. We searched MEDLINE via PubMed, the Cochrane Controlled Trials Register (Cochrane Library issue 01/2012) and Google Scholar in September 2013 to identify published reports, and the International Clinical Trials Registry Platform (ICTRP) of the World Health Organization (WHO) in March 2013 to identify registry records. Two readers recorded the non-inferiority margin and other data using a standardized data abstraction form.

Results: The margin was identical in study protocol and publication in 43 (80%) of 54 pairs of study protocols and articles. In the remaining pairs reporting was inconsistent (5 pairs, 9%), or the non-inferiority margin was either not reported in the publication (5 pairs, 9%) or not defined in the study protocol (1 pair). The confidence interval or the exact P-value required to judge whether the result was compatible with non-inferior, inferior or superior efficacy was reported in 43 (80%) publications. Complete and consistent reporting of both non-inferiority margin and confidence interval (or exact P-value) was present in 39 (72%) protocol-publication pairs. Twenty-nine trials (54%) were registered in trial registries, but only one registry record included the non-inferiority margin.

Conclusion: The reporting of non-inferiority margins was incomplete and inconsistent with study protocols in a substantial proportion of published trials, and margins were rarely reported in trial registries.

1. Introduction

A non-inferiority trial (NIT) measures a new treatment against a standard treatment to determine if it is not substantially worse. NITs are useful when benefits of standard therapy are known, and when novel treatments may be easier to use, less costly, or have fewer side effects [1]. NITs can also test pharmacologically related compounds to see if they are similarly effective [2]. A new treatment is considered non-inferior if the trial demonstrates that the new treatment is unlikely to be worse than an established treatment by more than a pre-specified amount, the *non-inferiority margin*. A non-inferiority margin that is too wide may compromise the results, and encourage acceptance and use of less effective therapies [3,4]. The number of published non-inferiority studies has substantially increased in recent years [5].

The interpretation of results of non-inferiority trials is challenging [4]. It requires an assessment of the rationale for the design and the assumptions underlying the choice of the non-inferiority margin [6]. Since readers have generally no access to study protocols the complete and accurate reporting of what was planned is essential [7]. Guidelines for the design and conduct of NITs have been issued by the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH)[8] and by the Committee for Medicinal Products for Human Use (CHMP) [9]. The Consolidated Standards of Reporting Trials (CONSORT) statement has been extended to improve the reporting of such trials [7]. The integrity of the NIT cannot be affirmed if authors do not accurately report the pre-specified non-inferiority margins and the relevant confidence intervals [10]. Authors must document the margins selected during the planning phase, and ensure that these margins are not chosen or modified *post hoc*, during analysis [4].

Some investigators modify design elements of a study, driven by their results. The *post hoc* modification of outcomes in randomized trials is a well-documented practice [11,12]. For example, a recent study of almost 3000 outcomes of (superiority) trials submitted to an ethics committee in Switzerland showed that in 30% of studies there were discrepancies between definitions in the protocols and in publications [13]. The risk of incorrect reporting is potentially greater for NITs than for superiority trials. In superiority trials the tested hypothesis is always the null hypothesis of no difference, which cannot be altered a posteriori. If the confidence interval on the difference includes 0, the new treatment is considered to be no better than the

reference treatment. In contrast, an NIT tests the hypothesis that the new treatment is less effective than the reference treatment by an acceptable amount, captured by the non-inferiority margin. Because the choice of the margin is to some extent arbitrary, researchers may be tempted to redefine the margin once the results are in, to claim non-inferiority. At present it is unknown whether this happens or not.

Our goal was to compare protocols of NITs submitted to ethics committees with published articles reporting the results of these NITs. We assessed the non-inferiority margins reported in protocols and publications, with the intent of determining whether the margins were concordant between protocols and publications. We also identified the studies that had been registered in a publicly accessible trial registry and examined whether or not the non-inferiority design and margin had been included in the registry record.

2. Methods

2.1 Identification of protocols of non-inferiority trials

In July 2012 we searched for protocols of non-inferiority trials in databases and archives of three research ethics committees: Kantonale Ethikkommission Bern (the Canton of Bern, Switzerland, see www.kek-bern.ch); Commission d'éthique de la recherche sur l'être humain (the Ethics Commission of University Hospitals of Geneva, Switzerland, see www.hug-ge.ch/ethique); and Ethische Commissie Leids Universitair Medisch Centrum (the ethics committee of Leiden University Medical Centre in the Netherlands, www.lumc.nl). We restricted our search to protocols submitted between January 1, 2001 and December 31, 2005. The non-inferiority design is relatively new, and few non-inferiority studies were published before 2001 [14]. We chose a cut-off date at the end of 2005 to allow enough time for the studies to be conducted and published.

A study protocol was eligible for inclusion if it described a non-inferiority or equivalence trial or stated that its goal was to determine whether a treatment was no worse than its comparator. When reviewing protocols for eligibility, we paid particular attention to the summary description of the study, the hypothesis that was tested, the statistical methods, and the determination of sample size. We included all non-inferiority trials, without regard to the number of arms, the intervention examined, or the inclusion of a non-inferiority margin in the protocol.

2.2 Identification of matching publications

In September 2012 we systematically searched for subsequent publications of each included study protocol in PubMed (National Library of Medicine), the Cochrane Collaboration's CENTRAL register of controlled clinical trials[15] (Cochrane Library, issue 01/2012), and Google Scholar. The CENTRAL database includes trials published in journals not indexed in MEDLINE, EMBASE or other bibliographic databases and in languages other than English. We searched for publications that used the study name or acronym (if available), combined with the condition studied, the intervention evaluated in the study and the names of the applicants. We included full text articles published in a medical journal. If more than one publication resulted from the same non-inferiority trial, we assessed the main publication. No language restrictions were applied to any of the searches.

2.3 Identification of registry records

In March 2013, we searched for registry records of each included study protocol in publicly accessible trial registries. We used the registration number reported in the publication or searched by study title, interventions and outcome in the International Clinical Trials Registry Platform (ICTRP) of the World Health Organization (WHO). The ICTRP search platform covers over ten trial registries, including ClinicalTrials.gov.

2.4 Data collection and definitions

We used a pre-tested, standardized data extraction form for study protocols, trial registry records and publications. Data were extracted by one investigator in each center (AP, JB MC, or OMD) and cross-checked by a second (MC, ME, OMD, SA, or TP). Discrepancies were resolved by consensus between the investigators. We extracted the following information from the study protocols (and, if available, from trial registry records): applicants, number and location of study centers, primary outcome, non-inferiority margin for the primary outcome, type of outcome (relative risk, risk difference, difference in continuous variables), type of experimental and control intervention (drug, device, surgical technique), planned sample size, subject area and source of funding (industry or non-commercial). Industry funding was defined as any financial support or provision of study materials by the pharmaceutical industry. If amendments were submitted, we considered the information in the latest amendment to be the correct protocol-defined non-inferiority margin. From published reports we extracted the year of publication, journal, reported study design, verdict regarding the experimental treatment (non-inferior, inferior, superior, inconclusive), and the upper limit of the confidence interval used to determine non-inferiority or exact p value.

2.5 Data analysis

We examined whether the non-inferiority margin reported in the publication was identical to the margin defined in the study protocol. We classified pairs of non-inferiority margins as concordant or discordant; the latter group was further divided by the direction of the difference (margin in the publication was larger or smaller than the margin in the protocol). For discordant cases, we assessed whether the change in the margin between protocol and publication also changed the verdict to non-inferiority. We examined whether the confidence interval of the estimate of the treatment effect was reported for the primary outcome. Information about both confidence interval and non-inferiority margin are necessary to reproduce the authors' verdict on

non-inferiority. If no confidence interval was reported we looked for a statistical test of non-inferiority. Complete and consistent reporting was defined as adequate reporting of both non-inferiority margins and confidence interval or exact p-value. We examined whether the sample size, industry funding or year of publication were associated with complete and consistent reporting in a logistic regression model. Finally, we identified trials that were described as non-inferiority trials in the protocol but described differently (for example as a superiority trial) in the publications. Data were analyzed using Stata software (version 12, College Station, Texas, USA).

2.6 Ethical approval

The ethics committees participating in this study and the authorities responsible for data protection in Bern and Geneva, Switzerland and Leiden, The Netherlands approved access to study protocols and their use for this research project. In Bern the approval was given under the condition that authors of included studies are not identified.

3. Results

3.1 Identification of eligible protocols and publications

We identified 123 eligible protocols that had been submitted to one of the three participating ethics committees (Bern, Geneva and Leiden) between January 1, 2001 and December 31, 2005 ([Figure 1](#)). We removed duplicates and then searched for publications that matched one of 115 unique study protocols. From these, we identified 54 publications, published between 2004 and 2012. The median year of submission to the ethics committee was 2003 both for protocols with and protocols without publication.

3.2 Characteristics of protocols and published reports

[Table 1](#) summarizes the characteristics of the 54 study protocols and corresponding publications. Forty-one (76%) trials were described as NITs and 13 (24%) as equivalence trials. Most trials were from the fields of infectious (15 trials) or vascular diseases (6), hematology (5) and diabetes (5). Publications increased in number from one in 2004 to thirteen in 2008 and decreased thereafter to three in 2011, and six in 2012. The 54 included studies were published in 37 different journals. Most studies were multicenter trials (51; 94%) and included countries other than Switzerland or The Netherlands (47; 92%). Most studies had planned sample sizes of 600 patients or more (33; 61%); most tested drugs (44; 81%) and received funding from the pharmaceutical industry (44; 81%).

Forty trials (74%) concluded that new treatments were non-inferior or equivalent; ten (19%) reported that the experimental intervention was superior to the control intervention; three (6%) reported that results were inconclusive (i.e. the confidence interval for the difference in treatment effects included the non-inferiority margin); one trial reported the new treatment was inferior (i.e. the confidence interval for the difference in treatment effects was below the non-inferiority margin).

3.3 Completeness and consistency of reporting

The authors of 50 publications (93%) explicitly reported the study design as a non-inferiority (or equivalence) trial or as a combined superiority and non-inferiority trial. Two publications were reported as superiority trials; two publications did not mention the design. [Table 2](#) details the reporting of non-inferiority margins in study protocols and matching publications, and the reporting of the relevant confidence intervals in

publications. In 43 (80%) of 54 pairs of study protocols and articles, non-inferiority margins were identical in study protocols and publications. In the remaining pairs reporting was inconsistent (5 pairs, 9%), or the non-inferiority margin was either not reported in the publication (5 pairs; 9%) or not defined in the study protocol (1 pair). The confidence interval or exact p-value required to judge whether the result was compatible with non-inferior, inferior or superior efficacy of the experimental treatment was reported in 43 (80%) publications. Complete and consistent reporting of both non-inferiority margins and confidence interval (or exact p-value) was present in 39 (72%) protocol-publication pairs. [Table 3](#) shows the characteristics of the studies with inconsistent or incomplete reporting. In a logistic regression model complete and consistent reporting was not associated with sample size ($p=0.54$), industry funding ($p=0.82$) or year of publication ($p=0.50$).

3.4 Reassessment of inconsistent and incomplete results

We reassessed the conclusion of non-inferiority for the 11 protocol-publication pairs that reported the non-inferiority margin inconsistently (5 pairs) or incompletely (6 pairs). The margin was wider in the publication than in the protocol for four trials. The protocol had a larger margin than the publication for one trial. The five inconsistent trials concluded that results demonstrated non-inferiority (4 trials) or superiority (1 trial). We found that these conclusions remained unchanged, even when we reassessed them in the light of the margin defined in the protocol.

The non-inferiority margin was not reported in the publication for five protocol-publication pairs. Among these five, the confidence interval appeared in one publication. For that study, the upper limit of the confidence interval reported in the publication was within the non-inferiority margin specified in the protocol. The missing confidence intervals in the publications of the remaining four pairs prevented us from assessing results in the light of the non-inferiority margin reported in the protocol. The single study that did not define the non-inferiority margin in the protocol did report it in the publication, but gave no confidence intervals.

3.5 Registry records

Out of 54 study protocols with matching publications, we identified 29 trials (54%) in publicly accessible registries. Twenty-eight were from ClinicalTrials.gov. The non-inferiority design was reported in five of these registry records (17%). Only one record (3%) reported the non-inferiority margin, which was consistent

with the one reported in the published report. Most trials (28, 97%) were registered before publication, 23 trials (82%) within one year after the study had been approved by the ethics committee, and 5 trials (18%) within two to five years. One trial was entered in the registry after publication.

4. Discussion

We compared the non-inferiority margins reported in study protocols approved by ethics committees in Switzerland and The Netherlands with the margins reported in subsequent publications. We found that non-inferiority margins had changed from protocol to published report in 5 (9%) out of 54 trials, but these changes did not affect the conclusions on non-inferiority. In a further 10 trials (19%) reporting was incomplete and in these cases it was generally not possible to reproduce the authors' conclusions regarding non-inferiority of the experimental intervention. Although our study did not show that researchers willfully manipulate non-inferiority margins, the incomplete reporting of results also meant that this could not be ruled out in a substantial proportion of trials.

Only about half of trials had been registered in a trial registry. The International Committee of Medical Journal Editors (ICMJE) requires public, prospective registration of clinical trials that began on or after July 1, 2005 [16]. For the present project, we included trials initiated between 2001 and 2005. The low registration rate for the trials we included might result from the failure of investigators to comply with the ICMJE request to retrospectively register all trials initiated before July 1, 2005. Our study confirms the results of a previous study from our group that compared publication of non-inferiority studies to information publicly accessible in trial registries. The study showed only 35 of 87 registry records described the design of the study as a NIT and only one record reported the non-inferiority margin [10]. For the present project, we included trials initiated between 2001 and 2005. The low registration rate for the trials we included might result from the failure of investigators to comply with the ICMJE request to retrospectively register all trials initiated before July 1, 2005.

We identified publications from 54 of 116 study protocols (46%). Although we conducted an extensive literature search we may have missed some publications because of delays between acceptance and publication of the manuscript, or because the study protocol did not provide sufficient information (e.g., name of all investigators involved in the study, brand name of intervention) to identify the publications that resulted from the study. We did not contact investigators to inquire about publications that we may have missed. Based on the experience gained in a previous study of protocols of drug trials and subsequent publications, about seven additional publications might have been identified in this way [17]. In the previous study, 233 out of

451 study protocols resulted in at least one publication, for a publication rate of 52% [17]. Studies of research proposals submitted to research ethics committees in France [18] and Spain [19] reported lower rates of publication (38% and 31%, respectively) whereas an older study from the United States of America reported higher rates (above 60%) [20].

Our aim was to examine protocol-publication pairs for consistent and transparent reporting. We did not assess whether the study design was appropriate for answering the research question. Nor did we check if the non-inferiority margin was used to correctly calculate the sample size [21], or whether it was considered in the statistical analysis plan of the study protocol. The methodological and reporting quality of reports of non-inferiority publications has been studied previously. One review of non-inferiority trials found that approximately only one fifth of the non-inferiority and equivalent studies provided an adequate rationale for the non-inferiority margin [22]. A study of the reporting quality of non-inferiority trials published 2003 to 2004 [23] found that less than 20% of the studies reported a clinical consideration and a justification of the margin. In the latter study the non-inferiority margin was not reported in only few articles (in six of 162 articles [3.7%]). Similarly, only five (2.2%) out of 232 non-inferiority trials identified in a PubMed search in 2009 did not report the margin, but less than half of the trials reported the method to determine the margin [24]. In our study the margin was missing somewhat more frequently in the published articles (in five of 54 articles [9.3%]), despite the fact that many of these studies were published more recently, after guidance on the transparent reporting of such trials had become available [25]. The extension to the Consolidated Standards of Reporting Trials (CONSORT) for non-inferiority and equivalence trials was first published in 2006 [25] and updated in 2012 [7]. The CONSORT reporting guidelines stress the importance of specifying the margins of non-inferiority or equivalence and the rationale for their choice [7,25].

Another limitation of our study is the relatively small sample size of our study: it is possible that a larger study might identify predictors associated with consistent and complete reporting of non-inferiority margins and confidence interval or p-values. For example, in a previous study of 227 protocol-publication pairs of drug trials we found that discrepant reporting of outcomes was more likely for smaller trials than for larger trials [13]. An important strength of the present study is that we had unrestricted access to the databases and archive of the three Ethics Committees. We could thus include an unselected sample of non-inferiority

studies submitted between 2001 and 2005 to one of three ethics committees in two European countries. However, the published trials included in this study may not be representative of all non-inferiority trials published in the literature. Indeed, compared to our study a sample of 232 non-inferiority trials identified in a PubMed search in 2009 included a larger proportion of smaller trials, with intermediate endpoints, and more trials sponsored by independent investigators [24]. Our study was dominated by larger, international multicenter trials funded by the pharmaceutical industry.

Finally, although this study is the first to empirically compare the reporting of the non-inferiority margin between study protocols and the subsequent publications but it reflects the earlier phase of the rise of non-inferiority and equivalence trials and may no longer be applicable to the studies submitted to ethics committees and institutional review boards today. The study should be repeated in the future to monitor the reporting of non-inferiority margins in study protocols and corresponding publications.

In conclusion, our findings emphasize that adhering to guidelines for protocol preparation and reporting recommendations is essential to the design and reporting of non-inferiority trials. At planning stage, recommendations about statistical considerations and the choice of non-inferiority margin should guide the development of protocols [26] and ensure their completeness. The non-inferiority margin must be specified in a study's planning stage and documentation of the margin must be transparent and complete so that the results of the study can be interpreted. As argued previously [10], trial registries should modify their databases so that the non-inferiority design and in particular the non-inferiority margin can be recorded in the registry. At publishing stage adherence to the extension of the CONSORT reporting guidelines to non-inferiority trials [7] will improve reports of non-inferiority trials. Our results support the addition of an item explicitly asking authors to report "any changes to non-inferiority margin after the trial commenced, with reasons" to a future update of the reporting guidelines. Finally, at reviewing stage, reviewers and editors should examine research protocols and plans for statistical analysis and ask authors to make such documents publicly available, as recommended by the ICMJE [27].

Acknowledgments

We are grateful to the Käthe-Zingg-Schwichtenberg-Fonds of the Swiss Academy of Medical Sciences for supporting this project. We thank all staff at the three Ethics Committees for their collaboration and in particular Dr. Dorothy Pfiffner for her help in identifying eligible trials in Bern.

References

- [1] Kaul S, Diamond GA. Good enough: a primer on the analysis and interpretation of noninferiority trials. *Ann Intern Med* 2006;145:62–9.
- [2] Ware JH, Antman EM. Equivalence trials. *N Engl J Med* 1997;337:1159–61.
- [3] Garattini S, Bertele V. Non-inferiority trials are unethical because they disregard patients' interests. *Lancet* 2007;370:1875–7.
- [4] Gøtzsche PC. Lessons from and cautions about noninferiority and equivalence randomized trials. *JAMA J Am Med Assoc* 2006;295:1172–4.
- [5] Suda KJ, Hurley AM, McKibbin T, Motl Moroney SE. Publication of noninferiority clinical trials: changes over a 20-year interval. *Pharmacotherapy* 2011;31:833–9.
- [6] Greene WL, Concato J, Feinstein AR. Claims of equivalence in medical research: are they supported by the evidence? *Ann Intern Med* 2000;132:715–22.
- [7] Piaggio G, Elbourne DR, Pocock SJ, Evans SJW, Altman DG, CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA J Am Med Assoc* 2012;308:2594–604.
- [8] The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human. Statistical principles for clinical trials E9. *Fed Regist n.d.*;63:49583–98.
- [9] Committee for Medicinal Products for Human Use, Efficacy Working Party, Committee for Release for Consultation. Committee for Medicinal Products for Human Use (CHMP) guideline on the choice of the non-inferiority margin. *Stat Med* 2006;25:1628–38.
- [10] Dekkers OM, Soonawala D, Vandenbroucke JP, Egger M. Reporting of noninferiority trials was incomplete in trial registries. *J Clin Epidemiol* 2011;64:1034–8.
- [11] Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA J Am Med Assoc* 2004;291:2457–65.
- [12] Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database Syst Rev* 2011:MR000031.
- [13] Redmond S, von Elm E, Blümle A, Gengler M, Gsponer T, Egger M. Cohort study of trials submitted to ethics committee identified discrepant reporting of outcomes in publications. *J Clin Epidemiol* 2013;66:1367–75.
- [14] Soonawala D, Middelburg RA, Egger M, Vandenbroucke JP, Dekkers OM. Efficacy of experimental treatments compared with standard treatments in non-inferiority trials: a meta-analysis of randomized controlled trials. *Int J Epidemiol* 2010;39:1567–81.
- [15] Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Eval Health Prof* 2002;25:38–64.
- [16] DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA J Am Med Assoc* 2004;292:1363–4.
- [17] Von Elm E, Röllin A, Blümle A, Huwiler K, Witschi M, Egger M. Publication and non-publication of clinical trials: longitudinal study of applications submitted to a research ethics committee. *Swiss Med Wkly* 2008;138:197–203.
- [18] Decullier E, Lheritier V, Chapuis F. Fate of biomedical research protocols and publication bias in France: retrospective cohort study. *Br Med J* 2005;331:19.
- [19] Pich J, Carne X, Arnaiz JA, Gomez B, Trilla A, Rodes J. Role of a research ethics committee in follow-up and publication of results. *Lancet* 2003;361:1015–6.
- [20] Dickersin K, Min YL, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA J Am Med Assoc* 1992;267:374–8.
- [21] Julious SA. Sample sizes for clinical trials with normal data. *Stat Med* 2004;23:1921–86.
- [22] Lange S, Freitag G. Choice of delta: requirements and reality--results of a systematic review. *Biom J Biom Z* 2005;47:12–27; discussion 99–107.
- [23] Le Henanff A, Giraudeau B, Baron G, Ravaud P. Quality of reporting of noninferiority and equivalence randomized trials. *JAMA J Am Med Assoc* 2006;295:1147–51.

- [24] Wangge G, Klungel OH, Roes KC, de Boer A, Hoes AW, Knol MJ. Room for improvement in conducting and reporting non-inferiority randomized controlled trials on drugs: a systematic review. *PLoS One* 2010;5:e13550.
- [25] Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA J Am Med Assoc* 2006;295:1152–60.
- [26] Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.
- [27] International Committee of Medical Journal Editors. *Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals*. 2013.

Figure 1: Flow diagram for selection of included study protocols and subsequent publications

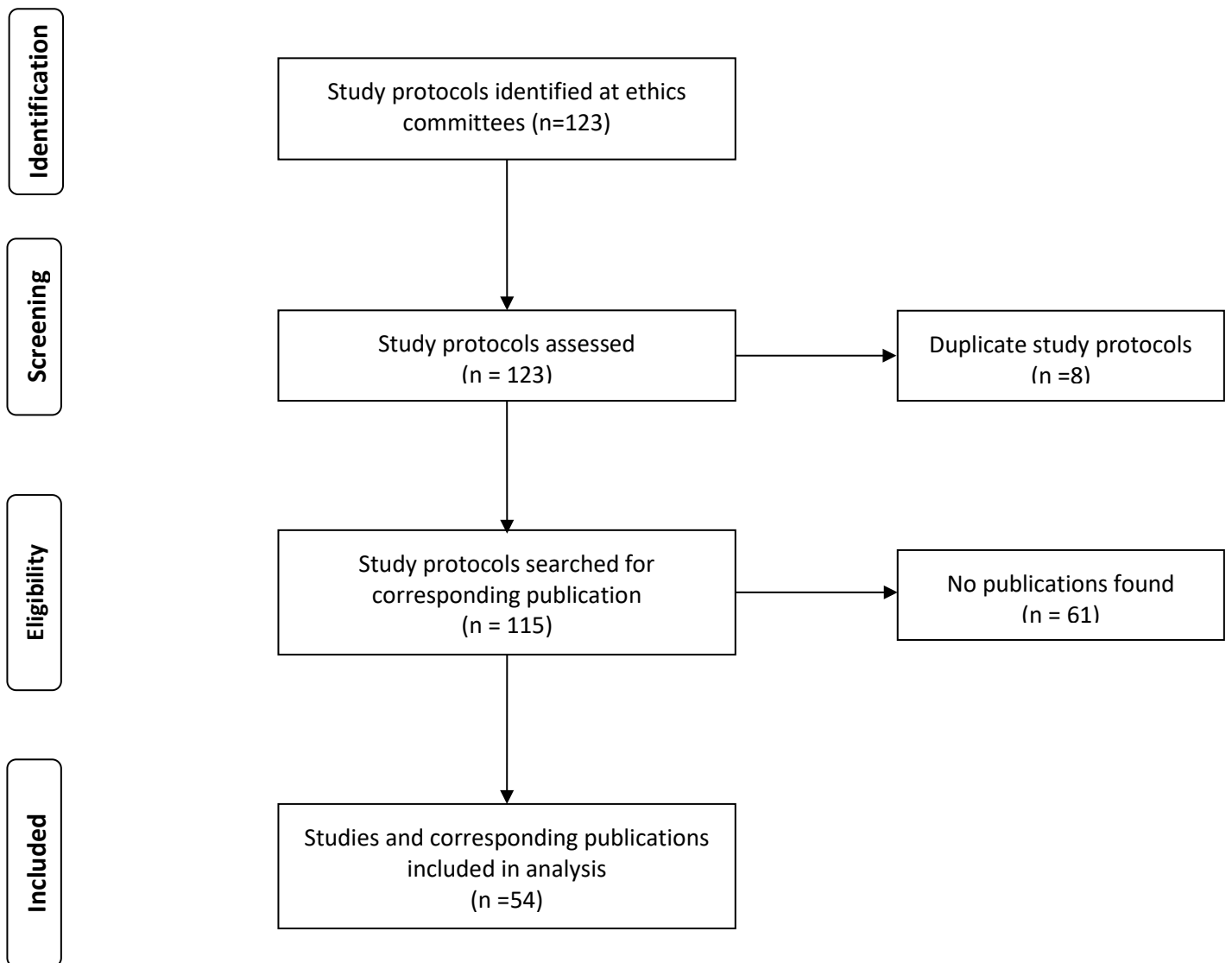


Table 1. Characteristics of 54 study protocols and matching publications of non-inferiority trials submitted to three ethics committees in Switzerland and The Netherlands 2001 to 2005.

| Study characteristics | No. (%) |
|--|----------------|
| Field of medicine | |
| Infectious diseases | 15 (28) |
| Vascular diseases | 6 (11) |
| Diabetology | 5 (9) |
| Hematology | 5 (9) |
| Rheumatology | 4 (7) |
| Other | 19 (35) |
| Publication year | |
| 2004 to 2006 | 15 (28) |
| 2007 to 2009 | 24 (44) |
| 2010 to 2012 | 15 (28) |
| Number of centers | |
| Multicenter | 51 (94) |
| Single center | 3 (6) |
| Source of funding | |
| Industry | 44 (81) |
| Other | 10 (19) |
| Type of outcome specified in protocol | |
| Binary outcome | 32 (59) |
| Continuous outcome | 15 (28) |
| Time-to event | 7 (13) |
| Planned sample size in protocol | |
| <400 | 11 (20) |
| 400 to <600 | 10 (19) |
| 600 to <1000 | 20 (37) |
| >1000 | 13 (24) |
| Type of intervention | |
| Drug | 44 (81) |
| Other | 10 (19) |
| Journal of publication | |
| Lancet | 9 (17) |
| New England Journal of Medicine | 3 (6) |
| American Journal of Transplantation | 2 (4) |
| Antimicrobial Agents and Chemotherapy | 2 (4) |
| Clinical Microbiology and Infection | 2 (4) |
| Current Medical Research Opinion | 2 (4) |
| Movement Disorders | 2 (4) |
| Pediatric Blood Cancer | 2 (4) |
| Other | 24 (44) |
| Verdict reported in the publication | |
| Non-inferior or equivalent | 40 (74) |
| Superior | 10 (19) |
| Inferiority not rejected | 3 (6) |
| Inferior | 1 (2) |

Table 2. Reporting of non-inferiority margins in 54 study protocols and matching publications and reporting of the relevant confidence intervals in publications.

| | No. (%) |
|---|----------------|
| Non-inferiority margin identical in protocol and publication | 43 (79) |
| Non-inferiority margin differs between protocol and publication | 5 (9) |
| Margin wider in publication than in protocol | 4 (7) |
| Margin wider in protocol than in publication | 1 (2) |
| Incomplete reporting of non-inferiority margin | 6 (11) |
| Margin reported in study protocol but not in publication | 5 (9) |
| Margin reported in publication but not in protocol | 1 (2) |
| Confidence interval or p-value for non-inferiority test | |
| Reported in publication | 43 (80) |
| Not reported in publication | 11 (20) |
| Complete and consistent reporting* | 39 (72) |

* Non-inferiority margins identical in protocol and publication and relevant confidence interval reported or p-value for the non-inferiority test reported in publication.

Table 3. Characteristics of 15 non-inferiority studies with inconsistent or incomplete reporting.

| Field | Year | Industry involved | Type of margin | Non-inferiority margin | | Upper limit of confidence interval | Authors' conclusion |
|--------------------------------------|------|-------------------|-----------------|------------------------|--------------|------------------------------------|---------------------|
| | | | | Protocol | Publication | | |
| <i>Inconsistent reporting</i> | | | | | | | |
| Infectious Diseases | 2004 | Yes | Risk difference | -0.12 | -0.10 | -0.012 | Non-inferior |
| Diabetology | 2005 | Yes | Risk difference | 0.13 | 0.15 | 0.007 | Non-inferior |
| Neurology | 2007 | Yes | Mean difference | 0.33 | 3 | -4.7 | Superior |
| Diabetology | 2010 | Yes | Mean difference | 0.3 | 0.4 | 0.2 | Non-inferior |
| Intensive care medicine | 2012 | Yes | Mean difference | -0.1 | -0.15 | -0.03 | Non-inferior |
| <i>Incomplete reporting</i> | | | | | | | |
| Infectious diseases | 2008 | No | Risk difference | Not reported | 0.04 | Not reported | Non-inferior |
| Cardiology | 2005 | Yes | Risk difference | 0.08 | Not reported | Not reported | Superior |
| Endocrinology | 2008 | Yes | Mean difference | -0.01 | Not reported | Not reported | Non-inferior |
| Diabetology | 2009 | Yes | Mean difference | 0.4 | Not reported | 0.1 | Non-inferior |
| Anesthesiology | 2009 | Yes | Risk difference | -0.05 | Not reported | Not reported | Non-inferior |
| Vascular Medicine | 2010 | No | Risk difference | 0.015 | Not reported | Not reported | Non-inferior |
| Ophthalmology | 2007 | Yes | Mean difference | 1 | 1 | Not reported | Non-inferior |
| Rheumatology | 2008 | Yes | Mean difference | -0.5 | -0.5 | Not reported | Non-inferior |
| Neurology | 2007 | Yes | Mean difference | -0.15 | -0.15 | Not reported | Inconclusive |
| Infectious diseases | 2007 | Yes | Risk difference | -0.1 | -0.1 | Not reported | Non-inferior |