BMJ Open

Testing for non-inferior mortality: a systematic review of non-inferiority margin sizes and trial characteristics

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-044480
Article Type:	Original research
Date Submitted by the Author:	03-Sep-2020
Complete List of Authors:	Pong, Sandra; The Hospital for Sick Children, Department of Pharmacy Urner, Martin; University of Toronto, Interdepartmental Division of Critical Care Medicine; University of Toronto, Institute of Health Policy, Management and Evaluation Fowler, Robert; Sunnybrook Health Sciences Centre, Tory Trauma Program; University of Toronto, Interdepartmental Division of Critical Care Medicine Mitsakakis, Nicholas ; University of Toronto, Dalla Lana School of Public Health Seto, Winnie; The Hospital for Sick Children, Department of Pharmacy; University of Toronto, Faculty of Pharmacy Hutchison, Jamie; The Hospital for Sick Children, Department of Critical Care Medicine Science, Michelle; The Hospital for Sick Children, Division of Infectious Diseases Daneman, Nick; Sunnybrook Health Sciences Centre, Division of Infectious Diseases, Department of Medicine
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, Clinical trials < THERAPEUTICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Testing for non-inferior mortality: a systematic review of non-inferiority margin sizes and trial characteristics

Sandra Pong (0000-0003-2963-9601) *clinical pharmacist*[1], Martin Urner *clinical and research fellow*[2-3], Robert A Fowler *clinician scientist*[2,4], Nicholas Mitsakakis *biostatistician*[5], Winnie Seto *clinical pharmacist*[1], James S Hutchison *senior scientist*[6], Michelle Science *infectious diseases consultant*[7], Nick Daneman *clinician scientist*[8]

[1]Department of Pharmacy, The Hospital for Sick Children, Toronto, Ontario, Canada
[2]Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto, Ontario Canada
[3]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
[4]Tory Trauma Program, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada
[5]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
[6]Department of Critical Care Medicine, The Hospital for Sick Children, Toronto, Ontario
[7]Division of Infectious Diseases, The Hospital for Sick Children, Toronto, Ontario, Canada
[8]Division of Infectious Diseases, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

Corresponding author:
Sandra Pong
The Hospital for Sick Children
555 University Avenue
Toronto, Ontario, M5G 1X8
Canada
sandra.pong@sickkids.ca

ABSTRACT

OBJECTIVE

To describe the size and variability of non-inferiority margins used in non-inferiority trials of medications with primary outcomes involving mortality, and to examine the association between trial characteristics and non-inferiority margin size.

DESIGN

Systematic review

DATA SOURCES

Medline, Medline In Process, Medline Epub Ahead of Print and Embase Classic+Embase databases from January 1989 to December 2019.

ELIGIBILITY CRITERIA

Prospective non-inferiority randomized controlled trials comparing pharmacological therapies, with primary analyses for non-inferiority and primary outcomes involving mortality alone or as part of a composite outcome. Trials had to pre-specify non-inferiority margins as absolute risk differences or relative to risks of outcome and provide a baseline risk of primary outcome in the control intervention.

RESULTS

3992 records were screened, 195 articles were selected for full text review and 111 articles were included for analyses. 82% of trials were conducted in thrombosis, infectious diseases or oncology. Mortality was the sole primary outcome in 23 (21%) trials, and part of a composite primary outcome in 88 (79%) trials. The overall median non-inferiority margin was an absolute risk difference of 9% (IQR 4.2-10%). When non-inferiority margins were expressed relative to the baseline risk of primary outcome in control groups, the median relative non-inferiority margin was 1.5 (IQR 1.3-1.7). In multivariable regression analyses examining the association between trial characteristics (medical specialty, inclusion of pediatric patients, mortality as a sole or part of a composite primary outcome, presence of industry funding) and non-inferiority margin size, only medical specialty was significantly associated with absolute non-inferiority margin size.

CONCLUSION

Absolute and relative non-inferiority margins used in published trials comparing medications are large, allowing conclusions of non-inferiority in the context of large differences in mortality. Accepting the potential for large increases in outcomes involving mortality while declaring non-inferiority is a challenging methodological issue in the conduct of non-inferiority trials.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- There have been no previous reviews or studies that describe the size and variability of non-inferiority margins used in trials with high-stake outcomes such as mortality.
- Our comprehensive and sensitive search for non-inferiority trials spanned a 30-year period to ensure that virtually all non-inferiority trials with primary outcome involving mortality would be captured.
- We were reliant on authors to provide the values of non-inferiority margins and estimated risks of outcomes in their sample size calculations.

WORD COUNT: 2840 words

INTRODUCTION

The premise of non-inferiority trials is to demonstrate that a new treatment is no worse than a standard intervention by a pre-specified non-inferiority margin chosen by researchers.[1] Yet proving that drugs, devices and other medical treatments are no worse than a comparison is challenging.[2-3] The acceptable width of the margin of non-inferiority is a controversial aspect in the design of these studies. It is a determinant of the required sample size of a trial and has a large influence on the interpretation of "not unacceptably worse." Wide margins allow smaller sample sizes to conclude non-inferiority, but if a margin is too wide, a conclusion of non-inferiority could be clinically irrelevant or ethically inappropriate. This would be especially disturbing if the implications of accepting a truly inadequate treatment as non-inferior involves death as an outcome.[2]

Design and analytical challenges, and the deficits in adherence to reporting standards of non-inferiority trials have been described in multiple studies and reviews.[4-12] Much attention has been focused on how non-inferiority margins are selected, whether they are justified[10,13] and how they affect the validity of trial results and conclusions.[11-12] However, prior research has not described the size and variability of non-inferiority margins used in trials with high-stake outcomes such as mortality, nor examined whether certain trial characteristics such as the type of patients, medical conditions studied, choice of outcomes and baseline risks of outcomes are associated with the selection of smaller or larger non-inferiority margins. There is a need to establish standards for the design and analyses of non-inferiority trials to promote consistent quality of these trials. An important step, therefore, is to identify the range of non-inferiority

margins used in non-inferiority trials and determine whether trial characteristics influence the selection of margin sizes.

In this systematic review, our primary objective was to describe the size and variability of non-inferiority margins used in non-inferiority trials of medications with primary outcomes involving mortality. Our secondary objective was to assess whether selected trial characteristics were associated with non-inferiority margin size.

METHODS

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement to report this systematic review.[14]

Search strategy

We searched Medline, Medline In Process, Medline Epub Ahead of Print and Embase Classic+Embase databases (OvidSP) (search performed February 8, 2019, updated December 12, 2019) to identify randomized controlled non-inferiority trials published between 1989 and 2019. Our decision to start our search from 1989 was informed by a review that described the changes in publication rate of non-inferiority trials between 1989 and 2009, and found 583 published non-inferiority trials but only one that was published prior to 1998.[3]

Subject heading and text-word terms for "equivalence trials or non-inferiority or inferiority studies" and mortality were used with the Cochrane sensitive trials filter. Of note, "non-

inferiority trial" and "inferiority trial" terms are indexed together with "equivalence trial" in Ovid and the term "equivalence trial" was only introduced as a Medical Subject Heading (MeSH) in 2018. Results were restricted to the English language and trials performed in humans. The complete electronic database search strategies are presented in Appendix A. To ensure that all relevant trials were captured, the electronic database search was supplemented with a manual search by scanning the reference lists of included trials and relevant reviews, in addition to a search of the reviewers' personal files.

Eligibility criteria

We included all prospective non-inferiority randomized controlled trials involving human subjects that compared pharmacological therapies, where the primary analysis was for non-inferiority and the primary outcome included mortality, either alone or as part of a composite outcome. All trials had to pre-specify a non-inferiority margin (as an absolute risk difference or relative to the risk of outcome) and provide a baseline estimate of the risk of primary outcome in the control intervention in a sample size calculation. In cases where these variables changed during the course of the trial, the initial values used in the original trial design were used for analyses. No distinction between pediatric or adult populations was made.

We excluded trials that did not provide a sample size calculation based on a pre-specified non-inferiority margin and estimated baseline risk of outcome. To enable comparisons of non-inferiority margins across different trials, we also excluded trials that used non-inferiority margins expressed as incidence rate ratios, odds ratios or hazard ratios because incidence and

hazard ratios are relative to an outcome event rate that changes with time and with odds ratios, the baseline risk of outcome in the control group cannot be determined to convert the ratio to a relative non-inferiority margin unless it was explicitly stated by the authors. We also excluded articles that described sub-studies, post-hoc analyses or follow-up studies of randomized trials.

Selection of trials

One review author (SP) screened titles and abstracts of all retrieved records for obvious exclusions. Two review authors (SP and MU) independently assessed potentially eligible trials based on full text review. Disagreements were resolved by arbitration by a third review author (ND).

Data collection

One review author (SP) extracted data from the included trials using a standardized form to collect information on: year of publication, medical specialty area, inclusion of pediatric patients (age less than 18 years), mortality as a single or part of a composite primary outcome, estimated risk of primary outcome in the control group, non-inferiority margin, industry funding (disclosures in the publication about funding or sponsorship by a pharmaceutical company) and conclusion about non-inferiority.

Statistical analyses

Trial characteristics were summarized using counts and proportions. To enable comparisons of non-inferiority margins across different trials as either absolute or relative margins, we converted non-inferiority margins expressed as absolute risk differences into relative non-inferiority margins relative to the estimated risk of outcome for each trial's control group. The reverse was also done to convert relative non-inferiority margins into equivalent margins in terms of absolute differences. Graphical plots were used to explore an association between absolute non-inferiority margins and the estimated risks of outcome in control groups, and to describe the distribution of absolute and relative non-inferiority margins used in the trials.

For the primary objective, descriptive statistics (median, interquartile range (IQR), range) of absolute and relative non-inferiority margins were summarized for the overall cohort of trials included in the review. We also stratified these by trial characteristics: medical specialty, inclusion of pediatric patients, mortality as a single or composite outcome, industry funding and publication date pre- or post-2010 release of the first FDA draft guidance statement about non-inferiority trials. To investigate whether there was a difference in non-inferiority margins (absolute and relative) according to trial characteristics, we compared non-inferiority margins using Wilcoxon rank sum test (for 2 groups) and Kruskal-Wallis rank sum test (for >2 groups).

For the secondary objective, we used multivariable linear regression to examine the association between pre-specified trial characteristics (medical specialty, inclusion of pediatric patients, mortality as single or composite outcome and industry funding) as independent variables and

non-inferiority margin size as the outcome variable. All comparisons were two-sided and p<0.05 was considered statistically significant. Statistical analyses were conducted using R version 4.0.2.

RESULTS

We screened 3992 records for relevance using titles and abstracts and selected 195 articles for full text review. After independent assessment of the full text articles and discussion among reviewers, a total of 111 articles met eligibility criteria to be included for analyses (Figure 1). The agreement between reviewers was excellent (kappa statistic = 0.86).

<<**Figure 1: PRISMA flow diagram**[14]>>

Trial characteristics

Among the 111 trials included, 91 (82%) were trials conducted in thrombosis, infectious diseases or oncology. Mortality was the sole primary outcome in 23 (21%) trials, and part of a composite primary outcome in 88 (79%) trials. Over half of the trials disclosed receiving some form of industry funding. Of the included trials, 82 (74%) concluded non-inferiority, 21 (19%) did not conclude non-inferiority and the remaining 8 (7%) were either inconclusive, stopped early or unclear about their conclusions. The non-inferiority margin was expressed as an absolute risk difference in 109 (98%) trials.

Association between absolute non-inferiority margins and estimated baseline risks of outcome (involving mortality) in control groups

Figure 2 is a scatterplot between absolute non-inferiority margins and estimated baseline risks of outcome (i.e. mortality alone or a composite outcome that included mortality) in the control group for the trials included in this review. A Spearman's correlation shows a moderate, positive monotonic correlation ($r_s = 0.6$, $p<0.05$) between the two. Variability in the absolute non-inferiority margins can be seen at both high and low estimates of baseline risks of outcome.

<<Figure 2: Association between absolute non-inferiority margins and estimated risks of outcome in control group>>

Distribution of non-inferiority margins for outcomes involving mortality

The distribution of absolute non-inferiority margins subdivided by medical specialty is shown in Figure 3A. There was a wide range of non-inferiority margins for trial outcomes that involve mortality (0.4 to 30%), with a skewed distribution and distinct peaks observed at 5, 10 and 15%. Thrombosis trials used smaller non-inferiority margins more commonly than did other trials.

Figure 3B illustrates a similarly skewed distribution of relative non-inferiority margins subdivided by medical specialty. The most common relative non-inferiority margin observed was 1.5. Most relative non-inferiority margins clustered in the range of 1.3 to 1.7, however there were also many relative non-inferiority margins that were greater than 2.

<<Figure 3: Distribution of absolute and relative non-inferiority margins for primary outcomes involving mortality>>

Characteristics of non-inferiority margins

The characteristics of the non-inferiority margins in the trials included in this review are summarized in Table 1. The median absolute non-inferiority margin was 9% (IQR 4.2-10%) and the median relative non-inferiority margin was 1.5 (IQR 1.3-1.7).

Table 1: Summary of characteristics of non-inferiority trials included

		Absolute non-inferiority margin (%) for outcomes involving mortality			Relative non-inferiority margin for outcomes involving mortality		
	n (%)	Median (IQR[a])	Range	p-value	Median (IQR[a])	Range	p-value
Overall	111	9 (4.2-10)	0.4-30	--	1.5 (1.3-1.7)	1.1-4.5	--
Medical specialty							
Thrombosis	37 (33.3)	3.6 (2-5)	0.4-30	<0.001	1.4 (1.3-1.7)	1.1-3.9	0.02
Infectious diseases	31 (27.9)	10 (5.5-10)	3-20		1.5 (1.5-1.7)	1.2-4.5	
Oncology	23 (20.7)	10 (7.5-13.8)	4-17.5		1.5 (1.3-1.9)	1.1-3	
Transplant	11 (9.9)	10 (10-13.5)	9-20		1.7 (1.5-2.5)	1.4-4.3	
Cardiology	4 (3.6)	7.5 (5-10.5)	5-12		1.3 (1.2-1.3)	1.1-1.4	
Gastroenterology	2 (1.8)	12.5 (11.3-13.8)	10-15		1.9 (1.6-2.2)	1.3-2.5	
Respirology	2 (1.8)	12.3 (11.1-13.4)	10-14.5		1.3 (1.27-1.3)	1.25-1.32	
Anesthesia	1 (0.9)	10 (10-10)	10-10		1.4 (1.4-1.4)	1.4-1.4	
Pediatric patients included							
Yes	21 (18.9)	10 (5-10)	3.5-15	0.11	1.7 (1.5-2)	1.2-4.5	0.10
No	85 (76.6)	8 (4-10)	0.7-30		1.5 (1.3-1.7)	1.1-4.3	
Unclear/not explicitly stated	5 (4.5)	5.5 (2.2-5.6)	0.4-10		1.4 (1.1-1.5)	1.1-3.2	
Mortality outcome							
Single	23 (20.7)	10 (5-12.8)	0.4-15	0.24	1.3 (1.2-1.6)	1.1-2.5	0.03
Composite	88 (79.3)	7.8 (4-10)	0.8-30		1.5 (1.4-1.7)	1.1-4.5	
Industry funding							
Yes	61 (55)	5.6 (3.5-10)	0.4-30	0.01	1.5 (1.3-1.7)	1.1-4.5	0.41
No	42 (37.8)	10 (5-10)	0.8-15		1.5 (1.4-1.8)	1.1-4	
Unclear/not explicitly stated	8 (7.2)	10 (10-13.4)	5.5-15		1.4 (1.3-2.7)	1.3-4.3	
Pre- and post-2010 release of draft FDA guidance statement							
Pre-2010	35 (31.5)	7.9 (3.8-10)	0.4-15	0.24	1.4 (1.2-1.7)	1.1-4	0.02
Post-2010	76 (68.5)	9.5 (4.5-10)	0.8-30		1.5 (1.4-1.7)	1.2-4.5	

[a]Interquartile range

The differences in both absolute and relative non-inferiority margins used among medical specialties were significant. Thrombosis trials had the lowest median absolute non-inferiority

margin of 3.5%. Although there was a wide range of absolute and relative non-inferiority margins used across trials, the absolute non-inferiority margins of at least one trial in every specialty was 10% or greater.

Trials with mortality as part of a composite primary outcome had significantly higher relative non-inferiority margins compared to those with mortality as a single primary outcome. In contrast, when the non-inferiority margin was expressed as an absolute risk difference, there was no significant difference in the margins between type of mortality outcome. Industry-funded trials had a significantly lower median absolute non-inferiority margin compared to those without industry funding.

In this review, 35 (32%) trials were published before 2010 when the first draft FDA guidance statement about non-inferiority trials was published. The relative non-inferiority margin sizes were significantly larger in trials that were published after 2010. A similar trend was seen with the absolute non-inferiority margin sizes, but the difference was not statistically significant.

Association between trial characteristics and non-inferiority margin size

Table 2 shows the β coefficients with 95% confidence intervals (CI) from multivariable linear regression analyses examining the association between trial characteristics (independent variables) and absolute non-inferiority margin size (outcome variable). Table 3 presents the regression analyses with relative non-inferiority margin size. Of the trial characteristics included in multivariable analyses, only medical specialty was significantly associated with the

size of absolute non-inferiority margin—specifically, thrombosis trials had significantly smaller absolute non-inferiority margins compared to trials in infectious diseases (reference group). In contrast, there were no statistically significant relationships observed in multivariable analyses between trial characteristics and relative non-inferiority margin size.

Table 2: Absolute non-inferiority margin regression analyses

Predictor	Adjusted β coefficient*	95% CI	p-value
Specialty			
Oncology	1.01	-1.24 to 3.26	0.38
Cardiovascular	-1.85	-6.24 to 2.55	0.41
Thrombosis	-5.81	-8.26 to -3.37	<0.001
Transplant	1.93	-1.28 to 5.14	0.24
Other[#]	3.15	-0.86 to 7.15	0.13
Infectious diseases	1 (reference)	--	--
Pediatrics			
Yes	-1.31	-3.62 to 1.01	0.27
No	1 (reference)	--	--
Mortality outcome			
Single	-0.29	-2.36 to 1.78	0.78
Composite	1 (reference)	--	--
Industry funding			
Yes	0.88	-1.11 to 2.87	0.39
No	1 (reference)	--	--

*Omnibus F-test: 6.75 (8, 102), p<0.001

[#]Due to low number of trials, "Other" category combines trials in anesthesia, gastroenterology and respirology

Table 3: Relative non-inferiority margin regression analyses

Predictor	Adjusted β coefficient*	95% CI	p-value
Specialty			
Oncology	-0.18	-0.52 to 0.16	0.31
Cardiovascular	-0.40	-1.07 to 0.26	0.24
Thrombosis	0.02	-0.35 to 0.40	0.90
Transplant	0.46	-0.03 to 0.95	0.07
Other[#]	-0.27	-0.88 to 0.34	0.39
Infectious diseases	1 (reference)	--	--
Pediatrics			
Yes	0.21	-0.15 to 0.56	0.25
No	1 (reference)	--	--
Mortality outcome			
Single	-0.21	-0.53 to 0.10	0.19
Composite	1 (reference)	--	--
Industry funding			
Yes	-0.25	-0.56 to 0.05	0.11
No	1 (reference)	--	--

*Omnibus F-test: 1.96 (8, 102), p=0.06

[#]Due to low number of trials, "Other" category combines trials in anesthesia, gastroenterology and respirology

DISCUSSION

We conducted a systematic review of 111 non-inferiority trials that compared pharmacological therapies where mortality was included in the primary outcome. We found that the majority of non-inferiority trials focused on thrombosis, infectious diseases and oncology. There was a wide range of non-inferiority margins used in these trials, irrespective of whether they were expressed as a measure of absolute effect or when converted to a relative effect. Our results showed that in the design of at least half of the non-inferiority trials included in this review, all of which included mortality as part of their primary outcome, a "new" drug therapy could have an absolute increase of 9% or relative increase of 50% in mortality outcomes compared to controls and still be accepted as non-inferior. Accepting the potential for this increase in

mortality while declaring non-inferiority is a challenging methodological issue in the conduct of non-inferiority trials.

In our review, we also found that non-inferiority margins were more commonly expressed in terms of absolute risk differences than in relative terms. Whether to present absolute or relative non-inferiority margins is a source of debate in the design of non-inferiority trials.[15] There is no clear consensus on the selection of the most appropriate effect measure but it has been demonstrated that different ways of expressing effect measures could result in different conclusions within the same non-inferiority trial.[15-17] Since a relative non-inferiority margin accounts for the estimated baseline risk of outcome, it would be a more conservative choice over an absolute margin to conclude non-inferiority should the event rate in the control group be lower than expected.

We detected significant variations in absolute and relative non-inferiority margin size according to medical specialty which could be partially explained by differences in acuity of diseases, patient age and life expectancy.[18] We also found that industry-funded trials had significantly lower median absolute non-inferiority margins compared to those without industry funding, presumably related to greater financial resources and higher capacity to support larger trials that are necessary when smaller non-inferiority margins are used.

When we compared non-inferiority margin size between trials published before and after the release of the 2010 draft FDA guidance document on non-inferiority studies, we found that the

median non-inferiority margin in trials published after 2010 was increased rather than decreased. Perhaps future guidelines could generate reductions in non-inferiority margins used for randomized controlled trials involving mortality, if they recommend margins lower, rather than higher than the current median (<9% absolute mortality, <1.5 times relative mortality).

There is currently limited research in the pre-defined determinants of the size of non-inferiority margins used in non-inferiority trials. Gayet-Ageron et al. conducted a survey among trialists to assess the association of pre-defined trial factors and non-inferiority margins. They found that lower non-inferiority margins were associated with mortality as a primary outcome, low baseline risk and lower costs of new treatments. In contrast, population age group and difficulties with patient recruitment did not appear to affect the choice of margin.[19] Because of the nature of a survey study, these results were based on self-report by respondents and were not necessarily reflective of actual practice when non-inferiority trials are designed and conducted.

In our review of published non-inferiority trials of drug therapies that included mortality as part of their primary outcome, we examined for an association between non-inferiority margin size and medical specialty, inclusion of pediatric patients, mortality as a single or composite outcome and presence of industry funding. The only significant association found was between absolute non-inferiority margins and medical specialty, specifically with thrombosis trials which we had already shown to have the lowest median absolute non-inferiority margin compared to

other trials. In contrast, when non-inferiority margins were expressed in relative terms, no significant associations could be found, even with thrombosis trials.

Similar to Gayet-Agergon et al.'s[19] results, we found a significant correlation between the size of absolute non-inferiority margins and estimated baseline risks of outcomes in the control group. While this association was moderate (r_s=0.6), it suggests that larger absolute non-inferiority margins are used when estimated risks of outcome occurring in control groups are higher. As can be seen in Figure 2, this relationship appears most evident for baseline outcome risks up to approximately 20%, beyond which larger absolute non-inferiority margins are no longer associated with higher baseline risk of outcome.

A strength of our review is the comprehensive and sensitive search for non-inferiority trials which spanned a 30-year period to ensure that virtually all non-inferiority trials with a primary outcome involving mortality would be captured. There were no limits placed on the type of medical specialty or patient population as long as a trial compared mortality between pharmacological therapies. However, we relied on authors to provide the values of non-inferiority margins and estimated risks of outcome in their sample size calculations within the publication or in their supplementary materials. The accuracy of reporting these variables was taken at face value. To enable standardized comparisons of absolute and relative non-inferiority margins to be made consistently across all trials included in the review, we omitted non-inferiority trials that used hazard ratios, odds ratios and event rate measures that either

changed with time or would not allow us to determine the estimated risk of outcome in the control group required for analyses.

The absolute and relative non-inferiority margins used in published trials comparing medications are large, allowing conclusions of non-inferiority in the context of large differences in mortality, and highly variable. Most trials utilize non-inferiority margins based on an absolute risk difference, which has only a moderate association with baseline estimates of risk for outcomes. With increasing popularity of non-inferiority trials, clinicians and other users of the medical literature should pay close attention to the size of non-inferiority margins used in these trials and consider the influence of study design parameters and inherent trial characteristics when interpreting the results. A collaborative effort to develop standards for the design and analyses of future non-inferiority trials would be beneficial to the scientific community.

REFERENCES

1 Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006;295:1152-60.

2 Mauri L, D'Agostino RB. Challenges in the design and interpretation of noninferiority trials. *N Engl J Med* 2017;377:1357-67.

3 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.

4 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. doi:10.1371/journal.pone.0013550

5 Schiller P, Burchardi N, Niestroj M, Kieser M. Quality of reporting of clinical non-inferiority and equivalence randomised trials—update and extension. *Trials* 2012;13:214. https://doi.org/10.1186/1745-6215-13-214

6 Lange S, Freitag G. Choice of delta: requirements and reality—results of a systematic review. *Biom J* 2005;47:12-27.

7 Althunian TA, de Boer A, Klungel OH, Insani WN, Groenwold RH. Methods of defining the non-inferiority margin in randomized, double-blind controlled trials: a systematic review. *Trials* 2017;18:107. https://doi.org/10.1186/s13063-017-1859-x

8 Rehal S, Morris TP, Fielding K, Carpenter JR, Phillips PP. Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. *BMJ Open* 2016;6:e012594. doi: 10.1136/bmjopen-2016-012594

9 Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. *Trials* 2011;12:106. https://doi.org/10.1186/1745-6215-12-106

10 D'Agostino RB, Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues—the encounters of academic consultants in statistics. *Statist Med* 2003;22:169-86.

11 Aberegg SK, Hersh AM, Samore MH. Empirical consequences of current recommendations for the design and interpretation of noninferiority trials. *J Gen Intern Med* 2018;33:88-96.

12 Hersh AM, Walter RJ, Abberegg SK. Use of mortality as an endpoint in noninferiority trials may lead to ethically problematic conclusions. *J Gen Intern Med* 2019;34:618-23.

13 Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. Non-inferiority clinical trials to establish effectiveness—guidance for industry. Silver Spring, MD: Food and Drug Administration, November 2016. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/non-inferiority-clinical-trials

14 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.

15 Abulizi X, Flandre P. Choice of treatment-effect measures when noninferiority margins originally defined in absolute difference translated into relative difference influenced the results of clinical trials. *J Clin Epidemiol* 2018;96:63-72.

16 Head SJ, Kaul S, Bogers AJ, Kappetein AP. Non-inferiority study design: lessons to be learned from cardiovascular trials. *Eur Heart J* 2012;33:1318-24.

17 Althunian TA, de Boer A, Groenwold RHH, Klungel OH. Defining the noninferiority margin and analysing noninferiority: an overview. *Br J Clin Pharmacol* 2017;83:1636-42.

18 Gladstone BP, Vach W. Choice of non-inferiority (NI) margins does not protect against degradation of treatment effects on an average—an observational study of registered and published NI trials. *PLoS ONE* 2014;9:e103616. doi:10.1371/journal.pone.0103616

19 Gayet-Ageron A, Agoritsas T, Rudaz S, Courvoisier D, Perneger T. The choice of the noninferiority margin in clinical trials was driven by baseline risk, type of primary outcome, and benefits of new treatment. *J Clin Epidemiol* 2015;68:1144-51.

STATEMENTS

Author contributions

SP, ND and RF drafted the study protocol. SP designed the data collection form, screened titles and abstracts, performed statistical analyses and drafted the study manuscript. SP and MU reviewed full-text articles and performed data collection. ND, MU, RF, NM, WS, JH and MS critically reviewed the final manuscript.

Licence statement

Competing interests declaration

All authors have completed the ICMJE uniform disclosure form and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Ethics approval

Not required

Funding statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. SP is supported by a SickKids Clinician-Scientist Training Program Scholarship from The Hospital for Sick Children. MU is supported by a Vanier Canada Graduate Scholarship from the Canadian Institutes of Health Research.

Patient and public involvement statement

It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Data sharing statement

Data and statistical code are available on request to the corresponding author.

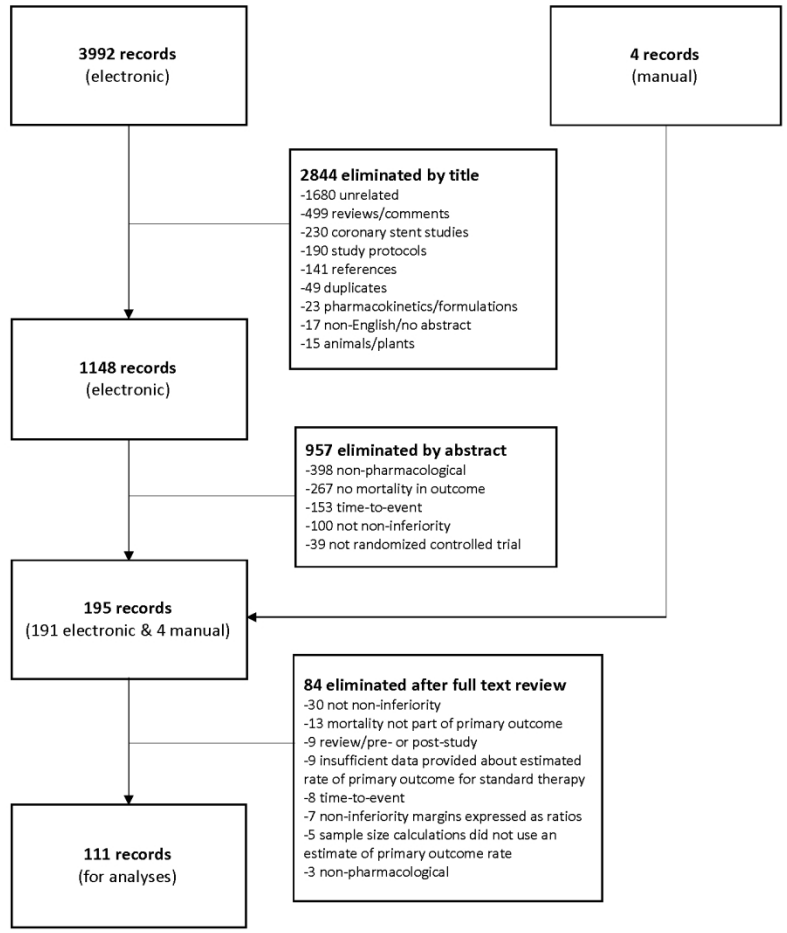


Figure 1: PRISMA flow diagram[14]

Figure 1: PRISMA flow diagram

215x279mm (200 x 200 DPI)

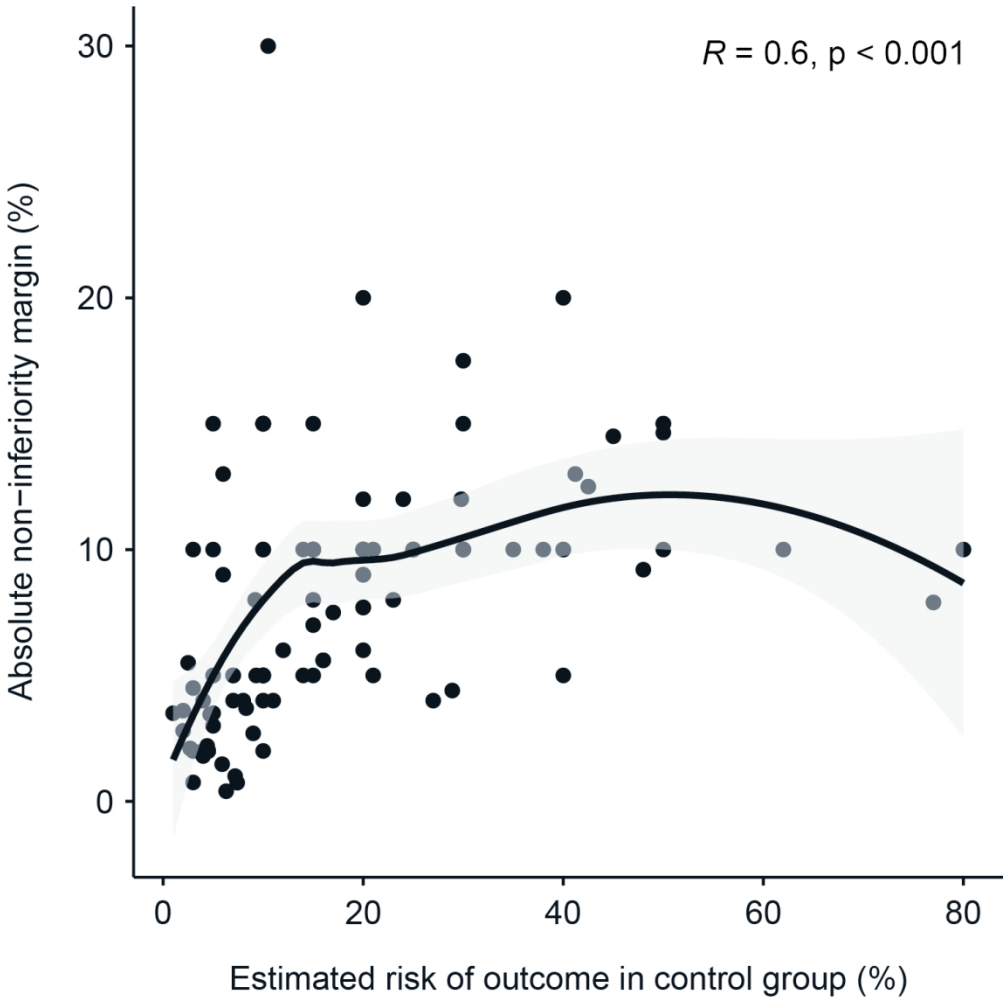


Figure 2: Association between absolute non-inferiority margins and estimated risks of outcome in control group

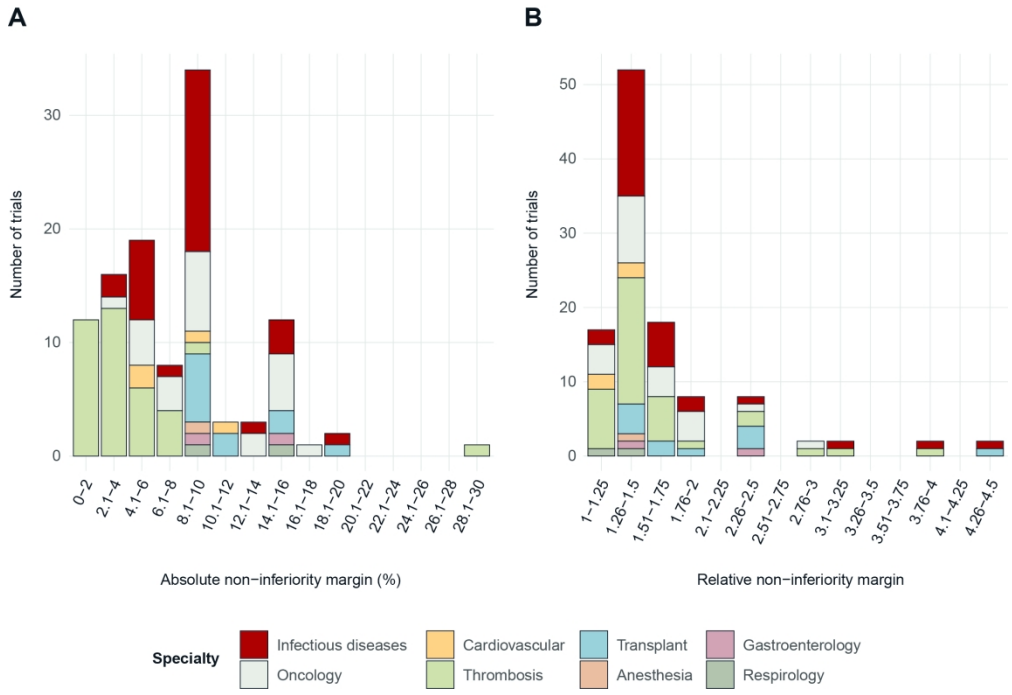


Figure 3: Distribution of absolute and relative non-inferiority margins for primary outcomes involving mortality

Appendix A: Search strategy

MEDLINE

	Searches	Results	Annotations
1	equivalence trial/	503	
2	equivalence trials as topic/	282	
3	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70165	
4	or/1-3	70414	
5	mo.fs.	561893	
6	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493002	
7	mortality/ or cause of death/ or fatal outcome/ or hospital mortality/ or infant mortality/ or maternal mortality/	213608	
8	or/5-7 [****mortality terms****]	1834719	
9	4 and 8 [****base clinical set****]	9115	
10	randomized controlled trial.pt.	500401	
11	controlled clinical trial.pt.	93574	
12	randomized.ab.	469732	
13	placebo.ab.	205181	
14	drug therapy.fs.	2180485	
15	randomly.ab.	327066	
16	trial.ab.	494306	
17	groups.ab.	2008069	
18	or/10-17	4632786	

19	exp animals/ not humans.sh.	4672546
20	18 not 19 [****Cochrane Handbook Highly Sensitive Search Strategy for identifying randomized trials (Box 6.4.c 2008 version)****]	4014138
21	9 and 20 [****Final results****]	3721
22	limit 21 to (english language and humans and yr="1989 -Current")	3212

Medline-in-Process

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	

Medline Epubs Ahead of Print

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	
13	equivalenc*adj3 trial*.ti,ab,kf.	0	
14	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
15	or/13-14	70227	
16	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
17	15 and 16 [****Base clinical set****]	5800	
18	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
19	placebo.ab.	205277	

20	randomly.ab.	327266
21	trial.ab.	494765
22	groups.ab.	2009271
23	or/18-22 [***Trial terms***]	2722197
24	17 and 23 [****Final results****]	2313

Embase Classic+Embase databases (OvidSP)

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	
13	equivalenc*adj3 trial*.ti,ab,kf.	0	

14	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227
15	or/13-14	70227
16	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992
17	15 and 16 [****Base clinical set****]	5800
18	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027
19	placebo.ab.	205277
20	randomly.ab.	327266
21	trial.ab.	494765
22	groups.ab.	2009271
23	or/18-22 [***Trial terms***]	2722197
24	17 and 23 [****Final results****]	2313
25	(equivalen* adj2 trial*).ti,ab.	758
26	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab.	69786
27	or/25-26 [***equivalency or non-inferiority terms****]	69985
28	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab.	1473665
29	mortality/ or cancer mortality/ or childhood mortality/ or embryo mortality/ or fetus mortality/ or infant mortality/ or maternal mortality/ or prenatal mortality/ or surgical mortality/ or perinatal mortality/ or newborn mortality/	77281
30	death/ or "cause of death"/ or dying/ or heart death/ or sudden death/ or child death/ or newborn death/	75385
31	or/28-30 [****mortality terms****]	1511412

32	27 and 31 [****Base clinical set****]	5823
33	randomized controlled trial/	500622
34	double-blind procedure/	0
35	single-blind procedure/	0
36	crossover-procedure/	0
37	random*.ti,ab,kw.	1109854
38	factorial*.ti,ab,kw.	29313
39	crossover*.ti,ab,kw.	62938
40	"cross over".ti,ab,kw.	22789
41	"cross-over*".ti,ab,kw.	23015
42	placebo*.ti,ab,kw.	212808
43	(doubl* adj5 blind*).ti,ab,kw.	149043
44	(singl* adj5 blind*).ti,ab,kw.	22640
45	assign*.ti,ab,kw.	308920
46	allocat*.ti,ab,kw.	115873
47	volunteer*.ti,ab,kw.	189496
48	or/33-47 [****Cochrane Box 6.3.2.2 EMBASE sensitive TherapyTreatment Effectiveness Filter terms****]	1731635
49	32 and 48 [***Cochrane trial filter****]	1802
50	ct.fs.	0
51	32 and 50 [***clinical trial subheading****]	0
52	limit 32 to (randomized controlled trial or controlled clinical trial or multicenter study)	1230

53	or/49,51-52 [****Final results****]	1974
54	limit 53 to conference abstract [Limit not valid in Ovid MEDLINE(R),Ovid MEDLINE(R) Daily Update,Ovid MEDLINE(R) In-Process,Ovid MEDLINE(R) Publisher; records were retained]	1974
55	53 not 54 [***Conference abstracts removed****]	0
56	limit 55 to (human and english language and yr="1989 -Current")	0

PRISMA 2009 Checklist

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

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	n/a
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	n/a
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	10
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	n/a
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	11-15
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	n/a
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	n/a
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	15-18
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	18-19
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	19
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	n/a

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Checklist

BMJ Open

Testing for non-inferior mortality: a systematic review of non-inferiority margin sizes and trial characteristics

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-044480.R1
Article Type:	Original research
Date Submitted by the Author:	03-Mar-2021
Complete List of Authors:	Pong, Sandra; The Hospital for Sick Children, Department of Pharmacy Urner, Martin; University of Toronto, Interdepartmental Division of Critical Care Medicine; University of Toronto, Institute of Health Policy, Management and Evaluation Fowler, Robert; Sunnybrook Health Sciences Centre, Tory Trauma Program; University of Toronto, Interdepartmental Division of Critical Care Medicine Mitsakakis, Nicholas ; University of Toronto, Dalla Lana School of Public Health Seto, Winnie; The Hospital for Sick Children, Department of Pharmacy; University of Toronto, Faculty of Pharmacy Hutchison, Jamie; The Hospital for Sick Children, Department of Critical Care Medicine Science, Michelle; The Hospital for Sick Children, Division of Infectious Diseases Daneman, Nick; Sunnybrook Health Sciences Centre, Division of Infectious Diseases, Department of Medicine
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Epidemiology
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, Clinical trials < THERAPEUTICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Testing for non-inferior mortality: a systematic review of non-inferiority margin sizes and trial characteristics

Sandra Pong (0000-0003-2963-9601) *clinical pharmacist*[1], Martin Urner *clinical and research fellow*[2-3], Robert A Fowler *clinician scientist*[2,4], Nicholas Mitsakakis *biostatistician*[5], Winnie Seto *clinical pharmacist*[1], James S Hutchison *senior scientist*[6], Michelle Science *infectious diseases consultant*[7], Nick Daneman *clinician scientist*[8]

[1]Department of Pharmacy, The Hospital for Sick Children, Toronto, Ontario, Canada
[2]Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto, Ontario Canada
[3]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
[4]Tory Trauma Program, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada
[5]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
[6]Department of Critical Care Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada
[7]Division of Infectious Diseases, The Hospital for Sick Children, Toronto, Ontario, Canada
[8]Division of Infectious Diseases, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

Corresponding author:
Sandra Pong
The Hospital for Sick Children
555 University Avenue
Toronto, Ontario, M5G 1X8
Canada
sandra.pong@sickkids.ca

ABSTRACT

OBJECTIVE

To describe the size and variability of non-inferiority margins used in non-inferiority trials of medications with primary outcomes involving mortality, and to examine the association between trial characteristics and non-inferiority margin size.

DESIGN

Systematic review

DATA SOURCES

Medline, Medline In Process, Medline Epub Ahead of Print and Embase Classic+Embase databases from January 1989 to December 2019.

ELIGIBILITY CRITERIA

Prospective non-inferiority randomized controlled trials comparing pharmacological therapies, with primary analyses for non-inferiority and primary outcomes involving mortality alone or as part of a composite outcome. Trials had to pre-specify non-inferiority margins as absolute risk differences or relative to risks of outcome and provide a baseline risk of primary outcome in the control intervention.

RESULTS

3992 records were screened, 195 articles were selected for full text review and 111 articles were included for analyses. 82% of trials were conducted in thrombosis, infectious diseases or oncology. Mortality was the sole primary outcome in 23 (21%) trials, and part of a composite primary outcome in 88 (79%) trials. The overall median non-inferiority margin was an absolute risk difference of 9% (IQR 4.2-10%). When non-inferiority margins were expressed relative to the baseline risk of primary outcome in control groups, the median relative non-inferiority margin was 1.5 (IQR 1.3-1.7). In multivariable regression analyses examining the association between trial characteristics (medical specialty, inclusion of pediatric patients, mortality as a sole or part of a composite primary outcome, presence of industry funding) and non-inferiority margin size, only medical specialty was significantly associated with non-inferiority margin size.

CONCLUSION

Absolute and relative non-inferiority margins used in published trials comparing medications are large, allowing conclusions of non-inferiority in the context of large differences in mortality. Accepting the potential for large increases in outcomes involving mortality while declaring non-inferiority is a challenging methodological issue in the conduct of non-inferiority trials.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- There have been no previous reviews or studies that describe the size and variability of non-inferiority margins used in trials with high-stake outcomes such as mortality.
- Our comprehensive and sensitive search for non-inferiority trials spanned a 30-year period to ensure that virtually all non-inferiority trials with primary outcome involving mortality would be captured.
- We were reliant on authors to provide the values of non-inferiority margins and estimated risks of outcomes in their sample size calculations.

WORD COUNT: 3003 words

INTRODUCTION

The premise of non-inferiority trials is to demonstrate that a new treatment is no worse than a standard intervention by a pre-specified non-inferiority margin chosen by researchers.[1] Yet proving that drugs, devices and other medical treatments are no worse than a comparison is challenging.[2-3] The acceptable width of the margin of non-inferiority is a controversial aspect in the design of these studies. It is a determinant of the required sample size of a trial and has a large influence on the interpretation of "not unacceptably worse." Wide margins allow smaller sample sizes to conclude non-inferiority, but if a margin is too wide, a conclusion of non-inferiority could be clinically irrelevant or ethically inappropriate. This would be especially disturbing if the implications of accepting a truly inadequate treatment as non-inferior involves death as an outcome.[2]

Design and analytical challenges, and the deficits in adherence to reporting standards of non-inferiority trials have been described in multiple studies and reviews.[4-12] Much attention has been focused on how non-inferiority margins are selected, whether they are justified[10,13] and how they affect the validity of trial results and conclusions.[11-12] The size of non-inferiority margins could also be influenced by the effectiveness of the standard treatment. A highly effective standard treatment could allow researchers to tolerate higher thresholds for decreased effectiveness with a new treatment.[14] However, prior research has not described the size and variability of non-inferiority margins used in trials with high-stake outcomes such as mortality, nor examined whether certain trial characteristics such as the type of patients, medical conditions studied, choice of outcomes and baseline risks of outcomes are associated

with the selection of smaller or larger non-inferiority margins. There is a need to establish standards for the design and analyses of non-inferiority trials to promote consistent quality of these trials. An important step, therefore, is to identify the range of non-inferiority margins used in non-inferiority trials and determine whether trial characteristics influence the selection of margin sizes.

In this systematic review, our primary objective was to describe the size and variability of non-inferiority margins used in non-inferiority trials of medications with primary outcomes involving mortality. Our secondary objective was to assess whether selected trial characteristics were associated with non-inferiority margin size. We hypothesized that non-inferiority margins in these trials will be large and variable; and the size of non-inferiority margins will be related to the type of patients and medical conditions studied, as well as availability of industry funding and how mortality has been included in the outcome.

METHODS

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement to report this systematic review.[15]

Search strategy

We searched Medline, Medline In Process, Medline Epub Ahead of Print and Embase Classic+Embase databases (OvidSP) (search performed February 8, 2019, updated December 12, 2019) to identify randomized controlled non-inferiority trials published between 1989 and

2019. Our decision to start our search from 1989 was informed by a review that described the changes in publication rate of non-inferiority trials between 1989 and 2009, and found 583 published non-inferiority trials but only one that was published prior to 1998.[3]

Subject heading and text-word terms for "equivalence trials or non-inferiority or inferiority studies" and mortality were used with the Cochrane sensitive trials filter. Of note, "non-inferiority trial" and "inferiority trial" terms are indexed together with "equivalence trial" in Ovid and the term "equivalence trial" was only introduced as a Medical Subject Heading (MeSH) in 2018. Results were restricted to the English language and trials performed in humans. The complete electronic database search strategies are presented in Appendix A. To ensure that all relevant trials were captured, the electronic database search was supplemented with a manual search by scanning the reference lists of included trials and relevant reviews, in addition to a search of the reviewers' personal files.

Eligibility criteria

We included all prospective non-inferiority randomized controlled trials involving human subjects that compared pharmacological therapies, where the primary analysis was for non-inferiority and the primary outcome included mortality, either alone or as part of a composite outcome. All trials had to pre-specify a non-inferiority margin (as an absolute risk difference or relative to the risk of outcome) and provide a baseline estimate of the risk of primary outcome in the control intervention in a sample size calculation. In cases where these variables changed

during the course of the trial, the initial values used in the original trial design were used for analyses. No distinction between pediatric or adult populations was made.

We excluded trials that did not provide a sample size calculation based on a pre-specified non-inferiority margin and estimated baseline risk of outcome. To enable comparisons of non-inferiority margins across different trials, we also excluded trials that used non-inferiority margins expressed as incidence rate ratios, odds ratios or hazard ratios because incidence and hazard ratios are relative to an outcome event rate that changes with time and with odds ratios, the baseline risk of outcome in the control group cannot be determined to convert the ratio to a relative non-inferiority margin unless it was explicitly stated by the authors. We also excluded articles that described sub-studies, post-hoc analyses or follow-up studies of randomized trials.

Selection of trials

One review author (SP) screened titles and abstracts of all retrieved records for obvious exclusions. Two review authors (SP and MU) independently assessed potentially eligible trials based on full text review. Disagreements were resolved by arbitration by a third review author (ND).

Data collection

One review author (SP) extracted data from the included trials using a standardized form to collect information on: year of publication, medical specialty area, inclusion of pediatric

patients (age less than 18 years), mortality as a single or part of a composite primary outcome, estimated risk of primary outcome in the control group, non-inferiority margin, industry funding (disclosures in the publication about funding or sponsorship by a pharmaceutical company) and conclusion about non-inferiority.

Statistical analyses

Trial characteristics were summarized using counts and proportions. To enable comparisons of non-inferiority margins across different trials as either absolute or relative margins, we converted non-inferiority margins expressed as absolute risk differences in percentages into relative non-inferiority margins relative to the estimated risk of outcome for each trial's control group. The reverse was also done to convert relative non-inferiority margins into equivalent margins in terms of absolute differences. Graphical plots were used to explore an association between absolute non-inferiority margins and the estimated risks of outcome in control groups, and to describe the distribution of absolute and relative non-inferiority margins used in the trials.

For the primary objective, descriptive statistics (median, interquartile range (IQR), range) of absolute and relative non-inferiority margins were summarized for the overall cohort of trials included in the review. We also stratified these by trial characteristics: medical specialty, inclusion of pediatric patients, mortality as a single or composite outcome, industry funding and publication date pre- or post-2010 release of the first FDA draft guidance statement about non-inferiority trials. To investigate whether there was a difference in non-inferiority margins

(absolute and relative) according to trial characteristics, we compared non-inferiority margins using Wilcoxon rank sum test (for 2 groups) and Kruskal-Wallis rank sum test (for >2 groups).

For the secondary objective, we used multivariable linear regression to examine the association between pre-specified trial characteristics (medical specialty, inclusion of pediatric patients, mortality as single or composite outcome and industry funding) as independent variables and non-inferiority margin size as the outcome variable. Due to the skewed distribution of the absolute and relative non-inferiority margins, we applied a log-transformation to the outcome variable which improved the performance and diagnostics of the regression models. All comparisons were two-sided and $p<0.05$ was considered statistically significant. Statistical analyses were conducted using R version 4.0.2.

RESULTS

We screened 3992 records for relevance using titles and abstracts and selected 195 articles for full text review. After independent assessment of the full text articles and discussion among reviewers, a total of 111 articles met eligibility criteria to be included for analyses (Figure 1). The agreement between reviewers was excellent (kappa statistic = 0.86).

<<**Figure 1: PRISMA flow diagram**[15]>>

Trial characteristics

Among the 111 trials included, 91 (82%) were trials conducted in thrombosis, infectious diseases or oncology. Mortality was the sole primary outcome in 23 (21%) trials, and part of a composite primary outcome in 88 (79%) trials. Over half of the trials disclosed receiving some form of industry funding. Of the included trials, 82 (74%) concluded non-inferiority, 21 (19%) did not conclude non-inferiority and the remaining 8 (7%) were either inconclusive, stopped early or unclear about their conclusions. The non-inferiority margin was expressed as an absolute risk difference in 109 (98%) trials. A summary of the included trials is provided in Appendix B.

Association between absolute non-inferiority margins and estimated baseline risks of outcome (involving mortality) in control groups

Figure 2 is a scatterplot between absolute non-inferiority margins and estimated baseline risks of outcome (i.e. mortality alone or a composite outcome that included mortality) in the control group for the trials included in this review. A Spearman's correlation shows a moderate, positive monotonic correlation ($r_s = 0.6$, $p<0.001$) between the two. Variability in the absolute non-inferiority margins can be seen at both high and low estimates of baseline risks of outcome. There was also a strong correlation between the observed outcomes reported in the trials and the initial estimated risks of outcome in the control groups ($r_s = 0.81$, $p<0.001$, Appendix C).

<<Figure 2: Association between absolute non-inferiority margins and estimated risks of outcome in control group>>

<<Appendix C: Relationship between observed outcomes and estimated risks of outcome in control group>>

Distribution of non-inferiority margins for outcomes involving mortality

The distribution of absolute non-inferiority margins subdivided by medical specialty is shown in Figure 3A. There was a wide range of non-inferiority margins for trial outcomes that involve mortality (0.4 to 30%), with a skewed distribution and distinct peaks observed at 5, 9 and 15%. Thrombosis trials used smaller non-inferiority margins more commonly than did other trials.

Figure 3B illustrates a similarly skewed distribution of relative non-inferiority margins subdivided by medical specialty. The most common relative non-inferiority margin observed was in the range of 1.26 to 1.5. Most relative non-inferiority margins clustered in the range of 1.3 to 1.7, however there were also many relative non-inferiority margins that were greater than 2.

<<**Figure 3: Distribution of absolute and relative non-inferiority margins for primary outcomes involving mortality**>>

Characteristics of non-inferiority margins

The characteristics of the non-inferiority margins in the trials included in this review are summarized in Table 1. The median absolute non-inferiority margin was 9% (IQR 4.2-10%) and the median relative non-inferiority margin was 1.5 (IQR 1.3-1.7).

Table 1: Summary of characteristics of non-inferiority trials included

		Absolute non-inferiority margin (%) for outcomes involving mortality			Relative non-inferiority margin for outcomes involving mortality		
	n (%)	Median (IQR[a])	Range	p-value	Median (IQR[a])	Range	p-value
Overall	111	9 (4.2-10)	0.4-30	--	1.5 (1.3-1.7)	1.1-4.5	--
Medical specialty							
Thrombosis	37 (33.3)	3.6 (2-5)	0.4-30	<0.001	1.4 (1.3-1.7)	1.1-3.9	0.02
Infectious diseases	31 (27.9)	10 (5.5-10)	3-20		1.5 (1.5-1.7)	1.2-4.5	
Oncology	23 (20.7)	10 (7.5-13.8)	4-17.5		1.5 (1.3-1.9)	1.1-3	
Transplant	11 (9.9)	10 (10-13.5)	9-20		1.7 (1.5-2.5)	1.4-4.3	
Cardiology	4 (3.6)	7.5 (5-10.5)	5-12		1.3 (1.2-1.3)	1.1-1.4	
Gastroenterology	2 (1.8)	12.5 (11.3-13.8)	10-15		1.9 (1.6-2.2)	1.3-2.5	
Respirology	2 (1.8)	12.3 (11.1-13.4)	10-14.5		1.3 (1.27-1.3)	1.25-1.32	
Anesthesia	1 (0.9)	10 (10-10)	10-10		1.4 (1.4-1.4)	1.4-1.4	
Pediatric patients included							
Yes	21 (18.9)	10 (5-10)	3.5-15	0.11	1.7 (1.5-2)	1.2-4.5	0.10
No	85 (76.6)	8 (4-10)	0.7-30		1.5 (1.3-1.7)	1.1-4.3	
Unclear/not explicitly stated	5 (4.5)	5.5 (2.2-5.6)	0.4-10		1.4 (1.1-1.5)	1.1-3.2	
Mortality outcome							
Single	23 (20.7)	10 (5-12.8)	0.4-15	0.24	1.3 (1.2-1.6)	1.1-2.5	0.03
Composite	88 (79.3)	7.8 (4-10)	0.8-30		1.5 (1.4-1.7)	1.1-4.5	
Industry funding							
Yes	61 (55)	5.6 (3.5-10)	0.4-30	0.01	1.5 (1.3-1.7)	1.1-4.5	0.41
No	42 (37.8)	10 (5-10)	0.8-15		1.5 (1.4-1.8)	1.1-4	
Unclear/not explicitly stated	8 (7.2)	10 (10-13.4)	5.5-15		1.4 (1.3-2.7)	1.3-4.3	
Pre- and post-2010 release of draft FDA guidance statement							
Pre-2010	35 (31.5)	7.9 (3.8-10)	0.4-15	0.24	1.4 (1.2-1.7)	1.1-4	0.02
Post-2010	76 (68.5)	9.5 (4.5-10)	0.8-30		1.5 (1.4-1.7)	1.2-4.5	

[a]Interquartile range

The differences in both absolute and relative non-inferiority margins used among medical specialties were significant. Thrombosis trials had the lowest median absolute non-inferiority margin of 3.6%. Although there was a wide range of absolute and relative non-inferiority margins used across trials, the absolute non-inferiority margins of at least one trial in every specialty was 10% or greater.

Trials with mortality as part of a composite primary outcome had significantly higher relative non-inferiority margins compared to those with mortality as a single primary outcome. In

contrast, when the non-inferiority margin was expressed as an absolute risk difference, there was no significant difference in the margins between type of mortality outcome. Industry-funded trials had a significantly lower median absolute non-inferiority margin compared to those without industry funding.

In this review, 35 (32%) trials were published before 2010 when the first draft FDA guidance statement about non-inferiority trials was published. The relative non-inferiority margin sizes were significantly larger in trials that were published after 2010. A similar trend was seen with the absolute non-inferiority margin sizes, but the difference was not statistically significant.

Association between trial characteristics and non-inferiority margin size

Table 2 shows the β coefficients with 95% confidence intervals (CI) from our regression analysis of the association between trial characteristics and the log-transformed absolute non-inferiority margin. Thrombosis trials had significantly smaller log-absolute non-inferiority margins compared to trials in infectious diseases (reference group) when adjusted for pediatric patients, single or composite mortality outcome and industry funding. When we analyzed the same for log-transformed relative non-inferiority margins, trials related to transplant had significantly larger log-relative non-inferiority margins compared to infectious diseases (Table 3).

Table 2: Absolute non-inferiority margin regression analyses

Predictor	Adjusted β coefficient*	95% CI	p-value
Specialty			
Oncology	0.12	-0.19 to 0.43	0.45
Cardiovascular	-0.23	-0.84 to 0.39	0.46
Thrombosis	-1.14	-1.48 to -0.8	<0.001
Transplant	0.15	-0.3 to 0.6	0.5
Other#	0.42	-0.13 to 0.98	0.14
Infectious diseases	1 (reference)	--	--
Pediatrics			
Yes	-0.22	-0.54 to 0.11	0.19
No	1 (reference)	--	--
Mortality outcome			
Single	-0.27	-0.56 to 0.02	0.07
Composite	1 (reference)	--	--
Industry funding			
Yes	0.08	-0.19 to 0.36	0.55
No	1 (reference)	--	--

*Omnibus F-test: 11.93 (8, 102), p<0.05, adjusted R-squared = 0.44

#Due to low number of trials, "Other" category combines trials in anesthesia, gastroenterology and respirology

Table 3: Relative non-inferiority margin regression analyses

Predictor	Adjusted β coefficient*	95% CI	p-value
Specialty			
Oncology	-0.08	-0.24 to 0.07	0.3
Cardiovascular	-0.23	-0.54 to 0.07	0.13
Thrombosis	0.01	-0.16 to 0.18	0.9
Transplant	0.24	0.01 to 0.46	0.04
Other[#]	-0.13	-0.41 to 0.15	0.35
Infectious diseases	1 (reference)	--	--
Pediatrics			
Yes	0.1	-0.06 to 0.26	0.23
No	1 (reference)	--	--
Mortality outcome			
Single	-0.11	-0.25 to 0.03	0.13
Composite	1 (reference)	--	--
Industry funding			
Yes	-0.13	-0.27 to 0.01	0.06
No	1 (reference)	--	--

*Omnibus F-test: 2.56 (8, 102), $p<0.05$, adjusted R-squared = 0.1

[#]Due to low number of trials, "Other" category combines trials in anesthesia, gastroenterology and respirology

DISCUSSION

We conducted a systematic review of 111 non-inferiority trials that compared pharmacological therapies where mortality was included in the primary outcome. We found that the majority of non-inferiority trials focused on thrombosis, infectious diseases and oncology. There was a wide range of non-inferiority margins used in these trials, irrespective of whether they were expressed as a measure of absolute effect or when converted to a relative effect. Our results showed that in the design of at least half of the non-inferiority trials included in this review, all of which included mortality as part of their primary outcome, a "new" drug therapy could have an absolute increase of 9% or relative increase of 50% in mortality outcomes compared to controls and still be accepted as non-inferior. Accepting the potential for this increase in

mortality while declaring non-inferiority is a challenging methodological issue in the conduct of non-inferiority trials.

In our review, we also found that non-inferiority margins were more commonly expressed in terms of absolute risk differences than in relative terms. Whether to present absolute or relative non-inferiority margins is a source of debate in the design of non-inferiority trials.[16] There is no clear consensus on the selection of the most appropriate effect measure but it has been demonstrated that different ways of expressing effect measures could result in different conclusions within the same non-inferiority trial.[14,16-17] Since a relative non-inferiority margin accounts for the estimated baseline risk of outcome, it would be a more conservative choice over an absolute margin to conclude non-inferiority should the event rate in the control group be lower than expected.

We detected significant variations in absolute and relative non-inferiority margin size according to medical specialty which could be partially explained by differences in acuity of diseases, patient age and life expectancy.[18] We also found that industry-funded trials had significantly lower median absolute non-inferiority margins compared to those without industry funding, presumably related to greater financial resources and higher capacity to support larger trials that are necessary when smaller non-inferiority margins are used. However, the difference was not significant when relative non-inferiority margins were compared between trials with and without industry funding.

When we compared non-inferiority margin size between trials published before and after the release of the 2010 draft FDA guidance document on non-inferiority studies, we found that the median non-inferiority margin in trials published after 2010 was increased rather than decreased. This was significant only for relative non-inferiority margins, but not for absolute non-inferiority margins. Perhaps future guidelines could generate reductions in non-inferiority margins used for randomized controlled trials involving mortality, if they recommend margins lower, rather than higher than the current median (<9% absolute mortality, <1.5 times relative mortality).

There is currently limited research in the pre-defined determinants of the size of non-inferiority margins used in non-inferiority trials. Gayet-Ageron et al. conducted a survey among trialists to assess the association of pre-defined trial factors and non-inferiority margins. They found that lower non-inferiority margins were associated with mortality as a primary outcome, low baseline risk and lower costs of new treatments. In contrast, population age group and difficulties with patient recruitment did not appear to affect the choice of margin.[19] Because of the nature of a survey study, these results were based on self-report by respondents and were not necessarily reflective of actual practice when non-inferiority trials are designed and conducted.

In our review of published non-inferiority trials of drug therapies that included mortality as part of their primary outcome, we examined for an association between non-inferiority margin size and medical specialty, inclusion of pediatric patients, mortality as a single or composite

outcome and presence of industry funding. Medical specialty was the only trial characteristic found to be significantly associated with the size of non-inferiority margins--specifically, thrombosis trials were associated with smaller absolute non-inferiority margins, while transplant trials were associated with larger relative non-inferiority margins.

Similar to Gayet-Agergon et al.'s[19] results, we found a significant correlation between the size of absolute non-inferiority margins and estimated baseline risks of outcomes in the control group. While this association was moderate ($r_s = 0.6$), it suggests that larger absolute non-inferiority margins are used when estimated risks of outcome occurring in control groups are higher. As can be seen in Figure 2, this relationship appears most evident for baseline outcome risks up to approximately 20%, beyond which larger absolute non-inferiority margins are no longer associated with higher baseline risk of outcome.

A strength of our review is the comprehensive and sensitive search for non-inferiority trials which spanned a 30-year period to ensure that virtually all non-inferiority trials with a primary outcome involving mortality would be captured. There were no limits placed on the type of medical specialty or patient population as long as a trial compared mortality between pharmacological therapies. However, we relied on authors to provide the values of non-inferiority margins and estimated risks of outcome in their sample size calculations within the publication or in their supplementary materials. The accuracy of reporting these variables was taken at face value. To enable standardized comparisons of absolute and relative non-inferiority margins to be made consistently across all trials included in the review, we omitted

non-inferiority trials that used hazard ratios, odds ratios and event rate measures that either changed with time or would not allow us to determine the estimated risk of outcome in the control group required for analyses.

The absolute and relative non-inferiority margins used in published trials comparing medications are large, allowing conclusions of non-inferiority in the context of large differences in mortality, and highly variable. Most trials utilize non-inferiority margins based on an absolute risk difference, which has only a moderate association with baseline estimates of risk for outcomes. With increasing popularity of non-inferiority trials, clinicians and other users of the medical literature should pay close attention to the size of non-inferiority margins used in these trials and consider the influence of study design parameters and inherent trial characteristics when interpreting the results. A collaborative effort to develop standards for the design and analyses of future non-inferiority trials would be beneficial to the scientific community.

REFERENCES

1 Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006;295:1152-60.

2 Mauri L, D'Agostino RB. Challenges in the design and interpretation of noninferiority trials. *N Engl J Med* 2017;377:1357-67.

3 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.

4 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. doi:10.1371/journal.pone.0013550

5 Schiller P, Burchardi N, Niestroj M, Kieser M. Quality of reporting of clinical non-inferiority and equivalence randomised trials—update and extension. *Trials* 2012;13:214. https://doi.org/10.1186/1745-6215-13-214

6 Lange S, Freitag G. Choice of delta: requirements and reality—results of a systematic review. *Biom J* 2005;47:12-27.

7 Althunian TA, de Boer A, Klungel OH, Insani WN, Groenwold RH. Methods of defining the non-inferiority margin in randomized, double-blind controlled trials: a systematic review. *Trials* 2017;18:107. https://doi.org/10.1186/s13063-017-1859-x

8 Rehal S, Morris TP, Fielding K, Carpenter JR, Phillips PP. Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. *BMJ Open* 2016;6:e012594. doi: 10.1136/bmjopen-2016-012594

9 Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. *Trials* 2011;12:106. https://doi.org/10.1186/1745-6215-12-106

10 D'Agostino RB, Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues—the encounters of academic consultants in statistics. *Statist Med* 2003;22:169-86.

11 Aberegg SK, Hersh AM, Samore MH. Empirical consequences of current recommendations for the design and interpretation of noninferiority trials. *J Gen Intern Med* 2018;33:88-96.

12 Hersh AM, Walter RJ, Abberegg SK. Use of mortality as an endpoint in noninferiority trials may lead to ethically problematic conclusions. *J Gen Intern Med* 2019;34:618-23.

13 Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. Non-inferiority clinical trials to establish effectiveness—guidance for industry. Silver Spring, MD: Food and Drug Administration, November 2016. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/non-inferiority-clinical-trials

14 Head SJ, Kaul S, Bogers AJ, Kappetein AP. Non-inferiority study design: lessons to be learned from cardiovascular trials. *Eur Heart J* 2012;33:1318-24.

15 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.

16 Abulizi X, Flandre P. Choice of treatment-effect measures when noninferiority margins originally defined in absolute difference translated into relative difference influenced the results of clinical trials. *J Clin Epidemiol* 2018;96:63-72.

17 Althunian TA, de Boer A, Groenwold RHH, Klungel OH. Defining the noninferiority margin and analysing noninferiority: an overview. *Br J Clin Pharmacol* 2017;83:1636-42.

18 Gladstone BP, Vach W. Choice of non-inferiority (NI) margins does not protect against degradation of treatment effects on an average—an observational study of registered and published NI trials. *PLoS ONE* 2014;9:e103616. doi:10.1371/journal.pone.0103616

19 Gayet-Ageron A, Agoritsas T, Rudaz S, Courvoisier D, Perneger T. The choice of the noninferiority margin in clinical trials was driven by baseline risk, type of primary outcome, and benefits of new treatment. *J Clin Epidemiol* 2015;68:1144-51.

STATEMENTS

Author contributions

SP, ND and RF drafted the study protocol. SP designed the data collection form, screened titles and abstracts, performed statistical analyses and drafted the study manuscript. SP and MU reviewed full-text articles and performed data collection. ND, MU, RF, NM, WS, JH and MS critically reviewed the final manuscript.

Licence statement

Competing interests declaration

All authors have completed the ICMJE uniform disclosure form and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Ethics approval

Not required

Funding statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. SP is supported by a SickKids Clinician-Scientist Training Program Scholarship from The Hospital for Sick Children. MU is supported by a Vanier Canada Graduate Scholarship from the Canadian Institutes of Health Research.

Patient and public involvement statement

It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Data sharing statement

Data and statistical code are available on request to the corresponding author.

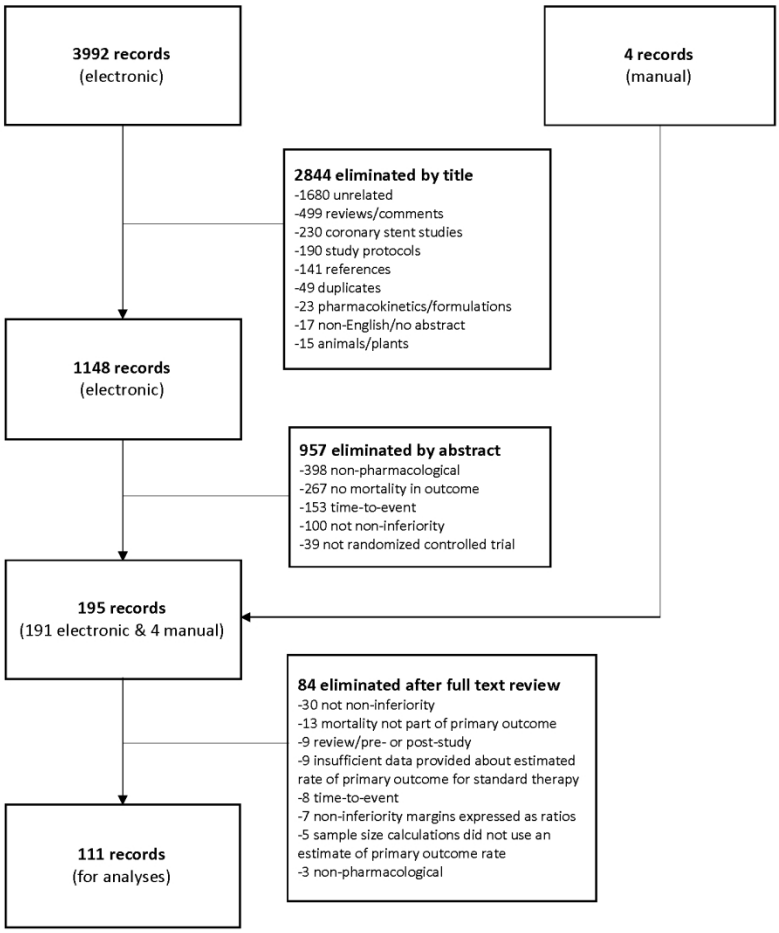


Figure 1: PRISMA flow diagram[14]

Figure 1: PRISMA flow diagram

215x279mm (200 x 200 DPI)

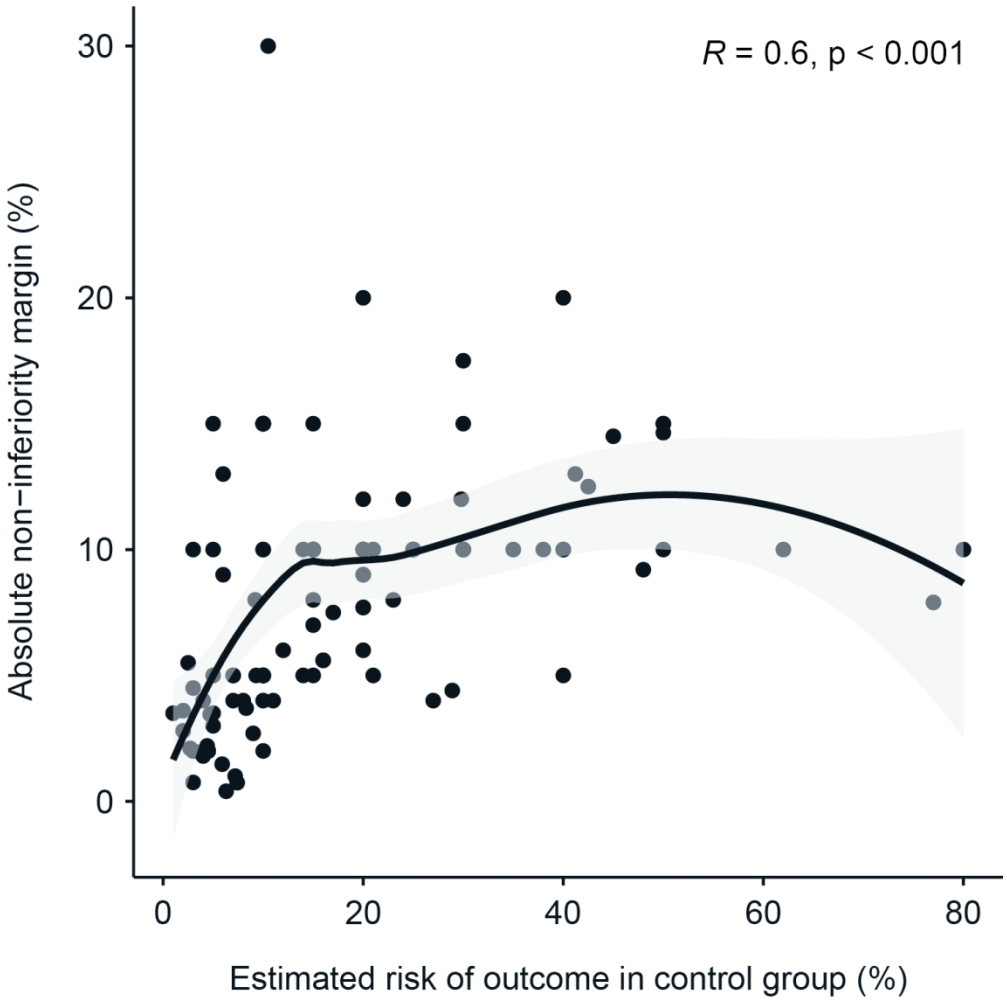


Figure 2: Association between absolute non-inferiority margins and estimated risks of outcome in control group

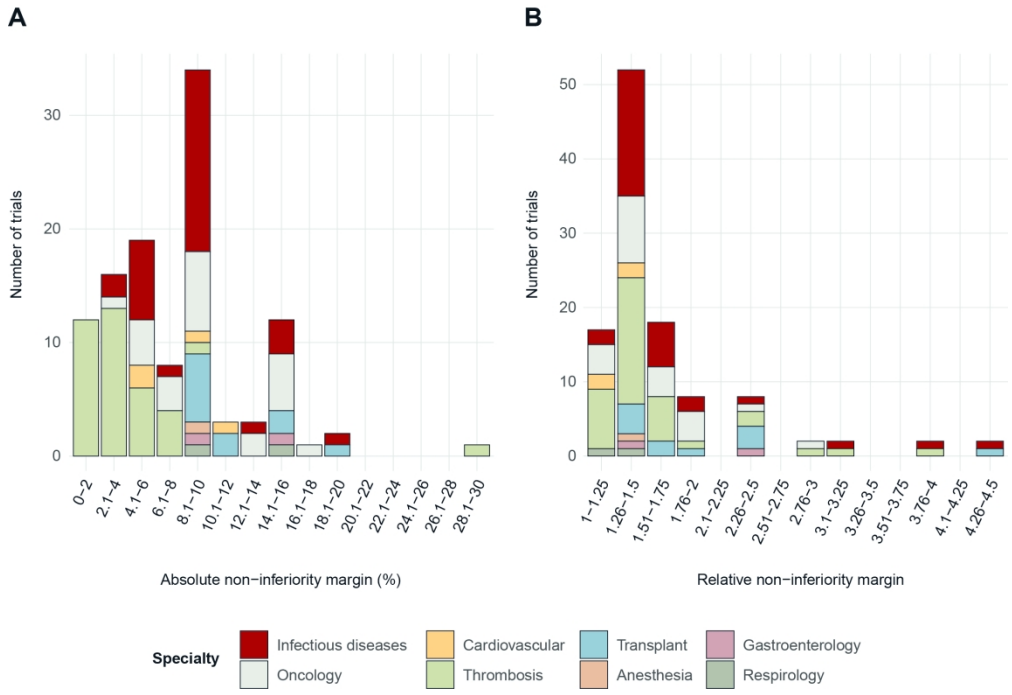


Figure 3: Distribution of absolute and relative non-inferiority margins for primary outcomes involving mortality

Appendix A: Search strategy

MEDLINE

	Searches	Results	Annotations
1	equivalence trial/	503	
2	equivalence trials as topic/	282	
3	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70165	
4	or/1-3	70414	
5	mo.fs.	561893	
6	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493002	
7	mortality/ or cause of death/ or fatal outcome/ or hospital mortality/ or infant mortality/ or maternal mortality/	213608	
8	or/5-7 [****mortality terms****]	1834719	
9	4 and 8 [****base clinical set****]	9115	
10	randomized controlled trial.pt.	500401	
11	controlled clinical trial.pt.	93574	
12	randomized.ab.	469732	
13	placebo.ab.	205181	
14	drug therapy.fs.	2180485	
15	randomly.ab.	327066	
16	trial.ab.	494306	
17	groups.ab.	2008069	
18	or/10-17	4632786	

19	exp animals/ not humans.sh.	4672546
20	18 not 19 [****Cochrane Handbook Highly Sensitive Search Strategy for identifying randomized trials (Box 6.4.c 2008 version)****]	4014138
21	9 and 20 [****Final results****]	3721
22	limit 21 to (english language and humans and yr="1989 -Current")	3212

Medline-in-Process

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	

Medline Epubs Ahead of Print

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	
13	equivalenc*adj3 trial*.ti,ab,kf.	0	
14	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
15	or/13-14	70227	
16	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
17	15 and 16 [****Base clinical set****]	5800	
18	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
19	placebo.ab.	205277	

20	randomly.ab.	327266
21	trial.ab.	494765
22	groups.ab.	2009271
23	or/18-22 [***Trial terms***]	2722197
24	17 and 23 [****Final results****]	2313

Embase Classic+Embase databases (OvidSP)

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	
13	equivalenc*adj3 trial*.ti,ab,kf.	0	

14	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227
15	or/13-14	70227
16	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992
17	15 and 16 [****Base clinical set****]	5800
18	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027
19	placebo.ab.	205277
20	randomly.ab.	327266
21	trial.ab.	494765
22	groups.ab.	2009271
23	or/18-22 [***Trial terms***]	2722197
24	17 and 23 [****Final results****]	2313
25	(equivalen* adj2 trial*).ti,ab.	758
26	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab.	69786
27	or/25-26 [***equivalency or non-inferiority terms****]	69985
28	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab.	1473665
29	mortality/ or cancer mortality/ or childhood mortality/ or embryo mortality/ or fetus mortality/ or infant mortality/ or maternal mortality/ or prenatal mortality/ or surgical mortality/ or perinatal mortality/ or newborn mortality/	77281
30	death/ or "cause of death"/ or dying/ or heart death/ or sudden death/ or child death/ or newborn death/	75385
31	or/28-30 [****mortality terms****]	1511412

32	27 and 31 [****Base clinical set****]	5823
33	randomized controlled trial/	500622
34	double-blind procedure/	0
35	single-blind procedure/	0
36	crossover-procedure/	0
37	random*.ti,ab,kw.	1109854
38	factorial*.ti,ab,kw.	29313
39	crossover*.ti,ab,kw.	62938
40	"cross over".ti,ab,kw.	22789
41	"cross-over*".ti,ab,kw.	23015
42	placebo*.ti,ab,kw.	212808
43	(doubl* adj5 blind*).ti,ab,kw.	149043
44	(singl* adj5 blind*).ti,ab,kw.	22640
45	assign*.ti,ab,kw.	308920
46	allocat*.ti,ab,kw.	115873
47	volunteer*.ti,ab,kw.	189496
48	or/33-47 [****Cochrane Box 6.3.2.2 EMBASE sensitive TherapyTreatment Effectiveness Filter terms****]	1731635
49	32 and 48 [***Cochrane trial filter****]	1802
50	ct.fs.	0
51	32 and 50 [***clinical trial subheading****]	0
52	limit 32 to (randomized controlled trial or controlled clinical trial or multicenter study)	1230

53	or/49,51-52 [****Final results****]	1974
54	limit 53 to conference abstract [Limit not valid in Ovid MEDLINE(R),Ovid MEDLINE(R) Daily Update,Ovid MEDLINE(R) In-Process,Ovid MEDLINE(R) Publisher; records were retained]	1974
55	53 not 54 [***Conference abstracts removed****]	0
56	limit 55 to (human and english language and yr="1989 -Current")	0

Appendix B: Non-inferiority trials

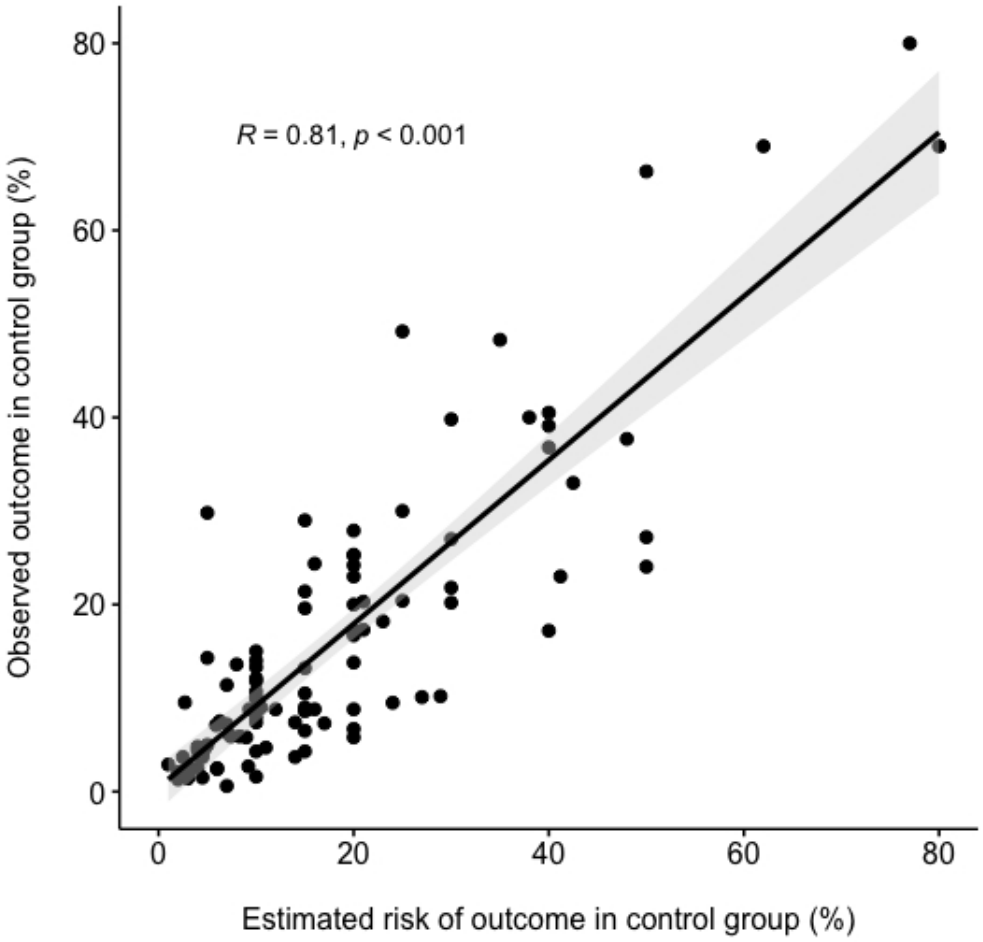
Author	Year	Country	Specialty	Pediatric patients included	Mortality as single or composite outcome	Industry funding	Conclusion	Estimated risk of event in control (%)	Observed event in control (%)	Absolute non-inferiority margin (% difference)	Relative non-inferiority margin
African Neonatal Sepsis Trial (AFRINEST) group	2015	international	infectious diseases	yes	composite	no	non-inferior	10	8	5	1.5
Ahn et al.	2013	Korea	oncology	no	single	not explicit/unclear	non-inferior	41.2	23	13	1.32
Aliberti et al.	2017	Italy	infectious diseases	no	composite	no	stopped early	10	7.4	5	1.5
Baqui et al.	2015	Bangladesh	infectious diseases	yes	composite	no	non-inferior	10	10	5	1.5
Bertrand et al.	2006	Canada	thrombosis	no	composite	yes	non-inferior	23	18.2	8	1.35
Beyer-Westendorf et al.	2017	Germany	thrombosis	no	composite	yes	non-inferior	3	1.7	4.5	2.5
Borchmann et al.	2017	international	oncology	no	composite	yes	non-inferior	12	8.8	6	1.5
Brack et al.	2012	international	infectious diseases	yes	composite	yes	not non-inferior	1	2.9	3.5	4.5
Budde et al.	2014	international	transplant	no	composite	yes	non-inferior	15	19.6	10	1.67
The Matisse Investigators	2003	international	thrombosis	no	composite	yes	non-inferior	5	5	3.5	1.7
Bunnapradist et al.	2013	international	transplant	no	composite	yes	non-inferior	6	2.5	9	2.5
Cai et al.	2014	China	transplant	no	composite	yes	non-inferior	20	16.7	20	2
Chastre et al.	2003	France	infectious diseases	no	single	no	non-inferior	40	17.2	10	1.25
Cordonnier et al.	2009	France	infectious diseases	no	single	no	non-inferior	9.2	2.7	8	1.87
de Kraker et al.	2004	international	oncology	yes	composite	no	non-inferior	15	8.6	10	1.67
De Simone et al.	2012	international	transplant	no	composite	yes	non-inferior	24	9.5	12	1.5
Diener et al.	2006	international	thrombosis	no	composite	yes	non-inferior	9.3	8.8	5	1.54
Eckardt et al.	2006	international	oncology	no	single	yes	non-inferior	62	69	10	1.16
Eriksson et al.	2011	international	thrombosis	no	composite	yes	non-inferior	20	8.8	7.7	1.39
Eriksson et al.	2007	international	thrombosis	no	composite	yes	non-inferior	48	37.7	9.2	1.19

Eriksson et al.	2007	international	thrombosis	no	composite	yes	non-inferior	20	6.7	7.7	1.39
Feres et al.	2013	Brazil	thrombosis	no	composite	yes	non-inferior	9	5.8	2.7	1.3
Ferme et al.	2017	international	oncology	yes	composite	no	non-inferior	10	10.1	10	2
Gaston et al.	2009	US	transplant	yes	composite	yes	non-inferior	20	27.9	10	1.5
Giamarellou et al.	2000	Greece	infectious diseases	no	composite	not explicit/unclear	non-inferior	25	49.2	10	1.4
Gilard et al.	2015	international	thrombosis	no	composite	yes	non-inferior	3	1.5	2	1.67
Gulizia et al.	2008	Italy	cardiology	no	composite	not explicit/unclear	non-inferior	38	40	10	1.26
Gwon et al.	2012	Korea	thrombosis	no	composite	yes	non-inferior	10	4.3	4	1.4
Sibbing et al.	2017	international	thrombosis	no	composite	yes	non-inferior	10.5	9	30	3.86
Han et al.	2016	China	thrombosis	no	composite	yes	non-inferior	8.3	5.9	3.7	1.45
Harris et al.	2018	international	infectious diseases	no	single	no	not non-inferior	14	3.7	5	1.36
Hofheinz et al.	2012	Germany	oncology	no	single	yes	non-inferior	42.5	33	12.5	1.29
Huh et al.	2017	Korea	transplant	no	composite	yes	non-inferior	10	13.3	15	2.5
Iversen et al.	2019	Denmark	infectious diseases	no	composite	no	non-inferior	10	12.1	10	2
Jeng et al.	2018	international	transplant	no	composite	yes	non-inferior	20	5.8	12	1.6
Johnson et al.	2016	international	oncology	no	composite	no	not non-inferior	5	14.3	5	2
Jones et al.	2016	Canada	anesthesia	no	composite	no	non-inferior	25	30	10	1.4
Kim et al.	2012	Korea	thrombosis	no	composite	yes	non-inferior	11	4.7	4	1.36
Park et al.	2013	Korea	thrombosis	no	composite	no	non-inferior	3	1.4	0.75	1.25
Hahn et al.	2018	Korea	thrombosis	no	composite	yes	non-inferior	4.5	4.2	2	1.44
Kirchhof et al.	2012	Germany	cardiology	no	composite	yes	not non-inferior	29.8	na	12	1.4
Kirchhof et al.	2018	international	thrombosis	no	composite	no	non-inferior	17	7.3	7.5	1.44
Kutner et al.	2015	US	cardiology	no	single	no	not non-inferior	21	20.3	5	1.24
Le et al.	2017	Vietnam	infectious diseases	no	single	no	non-inferior	15	6.5	10	1.67
Lee et al.	2018	Korea	thrombosis	no	composite	yes	non-inferior	7	0.6	4	1.57
Lee et al.	2009	UK	oncology	not explicit	single	no	non-inferior	80	69	10	1.13

Liu et al.	2013	China	oncology	yes	single	no	stopped early	15	na	15	2
Maertens et al.	2016	international	infectious diseases	no	single	yes	non-inferior	20	20	10	1.5
Mai et al.	2016	international	oncology	no	composite	yes	non-inferior	50	na	15	1.3
Matsumura-Nakano et al.	2019	Japan	thrombosis	no	composite	yes	not non-inferior	8	13.6	4	1.5
Mavroudis et al.	2016	Greece	oncology	no	composite	no	not non-inferior	15	10.5	7	1.47
Mavroudis et al.	2015	Greece	oncology	no	composite	no	not non-inferior	15	4.3	8	1.53
Merle et al.	2014	international	infectious diseases	no	composite	no	not non-inferior	20	17.2	6	1.3
Mesu et al.	2018	international	infectious diseases	yes	composite	no	non-inferior	6	2.4	13	3.17
Meynard et al.	2018	France	infectious diseases	no	composite	yes	not non-inferior	15	29	10	1.67
Mir et al.	2017	Pakistan	infectious diseases	yes	composite	no	non-inferior	10	11.8	5	1.5
Molloy et al.	2018	international	infectious diseases	no	single	no	non-inferior	15	21.4	10	1.67
Nakamura et al.	2017	Japan	thrombosis	no	composite	no	non-inferior	4.5	1.5	2	1.44
Nathan et al.	2005	Niger	infectious diseases	yes	composite	no	non-inferior	15	9	10	1.67
Park et al.	2014	Korea	gastroenterology	no	single	no	not non-inferior	10	na	15	2.5
Paul et al.	2015	Israel	infectious diseases	no	composite	no	not non-inferior	30	27	15	1.5
Ponticelli et al.	2014	Italy	transplant	no	composite	not explicit/unclear	inconclusive	3	2.8	10	4.33
Postma et al.	2015	Netherlands	infectious diseases	no	single	no	non-inferior	5	na	3	1.6
Pritchard-Jones et al.	2015	international	oncology	yes	composite	no	non-inferior	14	7.4	10	1.71
Pujade-Lauraine et al.	2010	international	oncology	no	composite	yes	non-inferior	77	80	7.9	1.1
Qazi et al.	2017	international	transplant	no	composite	yes	not non-inferior	25	20.4	10	1.4
Reynolds et al.	2010	Uganda	infectious diseases	no	composite	no	non-inferior	5	29.8	15	4
Riess et al.	2010	Germany	thrombosis	no	composite	yes	non-inferior	4.7	4.52	3.45	1.73
Russ et al.	2013	international	transplant	no	composite	yes	unclear	10	14	15	2.5
Schrappe et al.	2018	international	oncology	yes	composite	no	not non-inferior	4	4.4	4	2
Schroder et al.	2004	international	oncology	no	single	no	inconclusive	50	24.04	15	1.3

Schulz-Schupke et al.	2015	international	thrombosis	no	composite	yes	stopped early	10	1.6	2	1.2
Tedesco Silva et al.	2010	international	transplant	no	composite	yes	non-inferior	20	24.2	10	1.5
Sinha et al.	2005	international	respirology	yes	single	yes	non-inferior	45	na	14.5	1.32
Stabile et al.	2008	Italy	thrombosis	not explicit	composite	not explicit/unclear	non-inferior	2.5	3.7	5.5	3.2
Stellbrink et al.	2004	Germany	thrombosis	no	composite	yes	non-inferior	4	4.8	2	1.5
Stets et al.	2019	international	infectious diseases	no	composite	yes	non-inferior	21	17.3	10	1.48
Stone et al.	2007	international	thrombosis	no	composite	yes	not non-inferior	5.9	7.1	1.48	1.25
Swaminathan et al.	2011	India	infectious diseases	no	composite	no	not non-inferior	10	15	15	2.5
The GUSTO V Investigators	2001	international	thrombosis	no	single	yes	non-inferior	7.4	5.9	0.74	1.1
Turpie et al.	2009	international	thrombosis	no	composite	yes	non-inferior	27	10.1	4	1.15
Walsh et al.	2004	international	infectious diseases	yes	composite	yes	non-inferior	50	66.3	10	1.2
Willenheimer et al.	2005	international	cardiology	no	composite	yes	not non-inferior	40	36.8	5	1.13
Yahav et al.	2019	international	infectious diseases	no	composite	not explicit/unclear	non-inferior	35	48.3	10	1.29
Yakoub-Agha et al.	2012	international	oncology	no	single	not explicit/unclear	inconclusive	50	27.2	14.64	1.29
Yang et al.	2018	China	oncology	yes	composite	no	non-inferior	5	na	10	3
Daniels et al.	2019	South Africa	infectious diseases	yes	composite	no	non-inferior	7	11.4	5	1.71
Cisneros et al.	2019	international	infectious diseases	no	single	no	not non-inferior	20	25.3	10	1.5
Hahn et al.	2019	Korea	thrombosis	no	composite	yes	non-inferior	4	2.5	1.8	1.45
Jain et al.	2019	India	respirology	yes	composite	no	not non-inferior	40	39.1	10	1.25
Kollef et al.	2019	international	infectious diseases	no	single	yes	non-inferior	20	25.3	10	1.5
Nunn et al.	2019	international	infectious diseases	no	composite	no	non-inferior	30	20.2	10	1.33
Watanabe et al.	2019	Japan	thrombosis	not explicit	composite	yes	non-inferior	4.4	3.7	2.2	1.5
Charbonnier et al.	1998	international	thrombosis	no	composite	yes	non-inferior	7	7.2	5	1.71
Colombo et al.	2014	international	thrombosis	no	composite	yes	non-inferior	4.5	3.7	2	1.44
Lassen et al.	2010	international	thrombosis	not explicit	composite	yes	non-inferior	16	24.37	5.6	1.35

Lassen et al.	2009	international	thrombosis	no	composite	yes	not non-inferior	16	8.8	5.6	1.35
Le Deley et al.	2014	international	oncology	yes	composite	no	inconclusive	30	21.8	10	1.33
Noguchi et al.	2016	international	oncology	no	composite	yes	non-inferior	30	39.8	17.5	1.58
Patte et al.	1991	international	oncology	yes	composite	not explicit/unclear	non-inferior	10	10.7	15	2.5
Platzbecker et al.	2017	international	oncology	no	composite	no	non-inferior	15	13.2	5	1.33
Raffi et al.	2014	international	infectious diseases	no	composite	yes	non-inferior	20	13.8	9	1.45
Rubinstein et al.	2011	international	infectious diseases	no	composite	yes	non-inferior	40	40.5	20	1.5
Schulman et al.	2009	international	thrombosis	no	composite	yes	non-inferior	2	2.1	3.6	2.8
Schulman et al.	2013	international	thrombosis	no	composite	yes	non-inferior	2	1.3	2.8	2.4
Seo et al.	2014	Korea	gastroenterology	yes	composite	no	non-inferior	30	na	10	1.33
Continuous Infusion versus Double-Bolus Administration of Alteplase (COBALT) Investigators	1997	international	thrombosis	not explicit	single	yes	not non-inferior	6.3	7.53	0.4	1.06
Assessment of the Safety and Efficacy of a New Thrombolytic (ASSENT-2) Investigators	1999	international	thrombosis	no	single	yes	non-inferior	7.2	6.151	1	1.14
Vilas-Boas et al.	2014	Brazil	infectious diseases	yes	composite	no	non-inferior	20	23	9	1.45
International Joint Efficacy Comparison of Thrombolytics (INJECT)	1995	international	thrombosis	no	single	yes	non-inferior	2.7	9.53	2.1	1.78
Nitz et al.	2019	Germany	oncology	no	composite	yes	non-inferior	28.9	10.2	4.4	1.15



Appendix C: Relationship between observed outcomes and estimated risks of outcome in control group

188x180mm (72 x 72 DPI)

PRISMA 2009 Checklist

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

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	n/a
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	n/a
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	10
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	n/a
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	11-15
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	n/a
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	n/a
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	15-18
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	18-19
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	19
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	n/a

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Checklist

BMJ Open

Testing for non-inferior mortality: a systematic review of non-inferiority margin sizes and trial characteristics

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-044480.R2
Article Type:	Original research
Date Submitted by the Author:	30-Mar-2021
Complete List of Authors:	Pong, Sandra; The Hospital for Sick Children, Department of Pharmacy Urner, Martin; University of Toronto, Interdepartmental Division of Critical Care Medicine; University of Toronto, Institute of Health Policy, Management and Evaluation Fowler, Robert; Sunnybrook Health Sciences Centre, Tory Trauma Program; University of Toronto, Interdepartmental Division of Critical Care Medicine Mitsakakis, Nicholas ; University of Toronto, Dalla Lana School of Public Health Seto, Winnie; The Hospital for Sick Children, Department of Pharmacy; University of Toronto, Faculty of Pharmacy Hutchison, Jamie; The Hospital for Sick Children, Department of Critical Care Medicine Science, Michelle; The Hospital for Sick Children, Division of Infectious Diseases Daneman, Nick; Sunnybrook Health Sciences Centre, Division of Infectious Diseases, Department of Medicine
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Epidemiology
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, Clinical trials < THERAPEUTICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Testing for non-inferior mortality: a systematic review of non-inferiority margin sizes and trial characteristics

Sandra Pong (0000-0003-2963-9601) *clinical pharmacist*[1], Martin Urner *clinical and research fellow*[2-3], Robert A Fowler *clinician scientist*[2,4], Nicholas Mitsakakis *biostatistician*[5], Winnie Seto *clinical pharmacist*[1], James S Hutchison *senior scientist*[6], Michelle Science *infectious diseases consultant*[7], Nick Daneman *clinician scientist*[8]

[1]Department of Pharmacy, The Hospital for Sick Children, Toronto, Ontario, Canada
[2]Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto, Ontario Canada
[3]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
[4]Tory Trauma Program, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada
[5]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
[6]Department of Critical Care Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada
[7]Division of Infectious Diseases, The Hospital for Sick Children, Toronto, Ontario, Canada
[8]Division of Infectious Diseases, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

Corresponding author:
Sandra Pong
The Hospital for Sick Children
555 University Avenue
Toronto, Ontario, M5G 1X8
Canada
sandra.pong@sickkids.ca

ABSTRACT

OBJECTIVE

To describe the size and variability of non-inferiority margins used in non-inferiority trials of medications with primary outcomes involving mortality, and to examine the association between trial characteristics and non-inferiority margin size.

DESIGN

Systematic review

DATA SOURCES

Medline, Medline In Process, Medline Epub Ahead of Print and Embase Classic+Embase databases from January 1989 to December 2019.

ELIGIBILITY CRITERIA

Prospective non-inferiority randomized controlled trials comparing pharmacological therapies, with primary analyses for non-inferiority and primary outcomes involving mortality alone or as part of a composite outcome. Trials had to pre-specify non-inferiority margins as absolute risk differences or relative to risks of outcome and provide a baseline risk of primary outcome in the control intervention.

RESULTS

3992 records were screened, 195 articles were selected for full text review and 111 articles were included for analyses. 82% of trials were conducted in thrombosis, infectious diseases or oncology. Mortality was the sole primary outcome in 23 (21%) trials, and part of a composite primary outcome in 88 (79%) trials. The overall median non-inferiority margin was an absolute risk difference of 9% (IQR 4.2-10%). When non-inferiority margins were expressed relative to the baseline risk of primary outcome in control groups, the median relative non-inferiority margin was 1.5 (IQR 1.3-1.7). In multivariable regression analyses examining the association between trial characteristics (medical specialty, inclusion of pediatric patients, mortality as a sole or part of a composite primary outcome, presence of industry funding) and non-inferiority margin size, only medical specialty was significantly associated with non-inferiority margin size.

CONCLUSION

Absolute and relative non-inferiority margins used in published trials comparing medications are large, allowing conclusions of non-inferiority in the context of large differences in mortality. Accepting the potential for large increases in outcomes involving mortality while declaring non-inferiority is a challenging methodological issue in the conduct of non-inferiority trials.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- There have been no previous reviews or studies that describe the size and variability of non-inferiority margins used in trials with high-stake outcomes such as mortality.
- Our comprehensive and sensitive search for non-inferiority trials spanned a 30-year period to ensure that virtually all non-inferiority trials with primary outcome involving mortality would be captured.
- We were reliant on authors to provide the values of non-inferiority margins and estimated risks of outcomes in their sample size calculations.

WORD COUNT: 3110 words

INTRODUCTION

The premise of non-inferiority trials is to demonstrate that a new treatment is no worse than a standard intervention by a pre-specified non-inferiority margin chosen by researchers.[1] Yet proving that drugs, devices and other medical treatments are no worse than a comparison is challenging.[2-3] The acceptable width of the margin of non-inferiority is a controversial aspect in the design of these studies. It is a determinant of the required sample size of a trial and has a large influence on the interpretation of "not unacceptably worse." Wide margins allow smaller sample sizes to conclude non-inferiority, but if a margin is too wide, a conclusion of non-inferiority could be clinically irrelevant or ethically inappropriate. This would be especially disturbing if the implications of accepting a truly inadequate treatment as non-inferior involves death as an outcome.[2]

Design and analytical challenges, and the deficits in adherence to reporting standards of non-inferiority trials have been described in multiple studies and reviews.[4-12] Much attention has been focused on how non-inferiority margins are selected, whether they are justified[10,13] and how they affect the validity of trial results and conclusions.[11-12] The size of non-inferiority margins could also be influenced by the effectiveness of the standard treatment. A highly effective standard treatment could allow researchers to tolerate higher thresholds for decreased effectiveness with a new treatment.[14] However, prior research has not described the size and variability of non-inferiority margins used in trials with high-stake outcomes such as mortality, nor examined whether certain trial characteristics such as the type of patients, medical conditions studied, choice of outcomes and baseline risks of outcomes are associated

with the selection of smaller or larger non-inferiority margins. There is a need to establish standards for the design and analyses of non-inferiority trials to promote consistent quality of these trials. An important step, therefore, is to identify the range of non-inferiority margins used in non-inferiority trials and determine whether trial characteristics influence the selection of margin sizes.

In this systematic review, our primary objective was to describe the size and variability of non-inferiority margins used in non-inferiority trials of medications with primary outcomes involving mortality. Our secondary objective was to assess whether selected trial characteristics were associated with non-inferiority margin size. We hypothesized that non-inferiority margins in these trials will be large and variable; and the size of non-inferiority margins will be related to the type of patients and medical conditions studied, as well as availability of industry funding and how mortality has been included in the outcome.

METHODS

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement to report this systematic review.[15]

Search strategy

We searched Medline, Medline In Process, Medline Epub Ahead of Print and Embase Classic+Embase databases (OvidSP) (search performed February 8, 2019, updated December 12, 2019) to identify randomized controlled non-inferiority trials published between 1989 and

2019. Our decision to start our search from 1989 was informed by a review that described the changes in publication rate of non-inferiority trials between 1989 and 2009, and found 583 published non-inferiority trials but only one that was published prior to 1998.[3]

Subject heading and text-word terms for "equivalence trials or non-inferiority or inferiority studies" and mortality were used with the Cochrane sensitive trials filter. Of note, "non-inferiority trial" and "inferiority trial" terms are indexed together with "equivalence trial" in Ovid and the term "equivalence trial" was only introduced as a Medical Subject Heading (MeSH) in 2018. Results were restricted to the English language and trials performed in humans. The complete electronic database search strategies are presented in Appendix A. To ensure that all relevant trials were captured, the electronic database search was supplemented with a manual search by scanning the reference lists of included trials and relevant reviews, in addition to a search of the reviewers' personal files.

Eligibility criteria

We included all prospective non-inferiority randomized controlled trials involving human subjects that compared pharmacological therapies, where the primary analysis was for non-inferiority and the primary outcome included mortality, either alone or as part of a composite outcome. All trials had to pre-specify a non-inferiority margin (as an absolute risk difference or relative to the risk of outcome) and provide a baseline estimate of the risk of primary outcome in the control intervention in a sample size calculation. In cases where these variables changed

during the course of the trial, the initial values used in the original trial design were used for analyses. No distinction between pediatric or adult populations was made.

We excluded trials that did not provide a sample size calculation based on a pre-specified non-inferiority margin and estimated baseline risk of outcome. To enable comparisons of non-inferiority margins across different trials, we also excluded trials that used non-inferiority margins expressed as incidence rate ratios, odds ratios or hazard ratios because incidence and hazard ratios are relative to an outcome event rate that changes with time and with odds ratios, the baseline risk of outcome in the control group cannot be determined to convert the ratio to a relative non-inferiority margin unless it was explicitly stated by the authors. We also excluded articles that described sub-studies, post-hoc analyses or follow-up studies of randomized trials.

Selection of trials

One review author (SP) screened titles and abstracts of all retrieved records for obvious exclusions. Two review authors (SP and MU) independently assessed potentially eligible trials based on full text review. Disagreements were resolved by arbitration by a third review author (ND).

Data collection

One review author (SP) extracted data from the included trials using a standardized form to collect information on: year of publication, medical specialty area, inclusion of pediatric

patients (age less than 18 years), mortality as a single or part of a composite primary outcome, estimated risk of primary outcome in the control group, non-inferiority margin, industry funding (disclosures in the publication about funding or sponsorship by a pharmaceutical company) and conclusion about non-inferiority.

Statistical analyses

Trial characteristics were summarized using counts and proportions. To enable comparisons of non-inferiority margins across different trials as either absolute or relative margins, we converted non-inferiority margins expressed as absolute risk differences in percentages into relative non-inferiority margins relative to the estimated risk of outcome for each trial's control group. The reverse was also done to convert relative non-inferiority margins into equivalent margins in terms of absolute differences. Graphical plots were used to explore an association between absolute non-inferiority margins and the estimated risks of outcome in control groups, and to describe the distribution of absolute and relative non-inferiority margins used in the trials.

For the primary objective, descriptive statistics (median, interquartile range (IQR), range) of absolute and relative non-inferiority margins were summarized for the overall cohort of trials included in the review. We also stratified these by trial characteristics: medical specialty, inclusion of pediatric patients, mortality as a single or composite outcome, industry funding and publication date pre- or post-2010 release of the first FDA draft guidance statement about non-inferiority trials. To investigate whether there was a difference in non-inferiority margins

(absolute and relative) according to trial characteristics, we compared non-inferiority margins using Wilcoxon rank sum test (for 2 groups) and Kruskal-Wallis rank sum test (for >2 groups).

For the secondary objective, we used multivariable linear regression to examine the association between pre-specified trial characteristics (medical specialty, inclusion of pediatric patients, mortality as single or composite outcome and industry funding) as independent variables and non-inferiority margin size as the outcome variable. Due to the skewed distribution of the absolute and relative non-inferiority margins, we applied a log-transformation to the outcome variable to improve the performance and diagnostics of the regression models. All comparisons were two-sided and $p<0.05$ was considered statistically significant. Statistical analyses were conducted using R version 4.0.2.

RESULTS

We screened 3992 records for relevance using titles and abstracts and selected 195 articles for full text review. After independent assessment of the full text articles and discussion among reviewers, a total of 111 articles met eligibility criteria to be included for analyses (Figure 1). The agreement between reviewers was excellent (kappa statistic = 0.86).

<<**Figure 1: PRISMA flow diagram**[15]>>

Trial characteristics

Among the 111 trials included, 91 (82%) were trials conducted in thrombosis, infectious diseases or oncology. Mortality was the sole primary outcome in 23 (21%) trials, and part of a composite primary outcome in 88 (79%) trials. Over half of the trials disclosed receiving some form of industry funding. Of the included trials, 82 (74%) concluded non-inferiority, 21 (19%) did not conclude non-inferiority and the remaining 8 (7%) were either inconclusive, stopped early or unclear about their conclusions. The non-inferiority margin was expressed as an absolute risk difference in 109 (98%) trials. A summary of the included trials is provided in Appendix B.

Association between absolute non-inferiority margins and estimated baseline risks of outcome (involving mortality) in control groups

Figure 2 is a scatterplot between absolute non-inferiority margins and estimated baseline risks of outcome (i.e. mortality alone or a composite outcome that included mortality) in the control group for the trials included in this review. A Spearman's correlation shows a moderate, positive monotonic correlation ($r_s = 0.6$, $p<0.001$) between the two. Variability in the absolute non-inferiority margins can be seen at both high and low estimates of baseline risks of outcome. There was also a strong correlation between the observed outcomes reported in the trials and the initial estimated risks of outcome in the control groups ($r_s = 0.81$, $p<0.001$, Appendix C).

<<Figure 2: Association between absolute non-inferiority margins and estimated risks of outcome in control group>>

<<Appendix C: Relationship between observed outcomes and estimated risks of outcome in control group>>

Distribution of non-inferiority margins for outcomes involving mortality

The distribution of absolute non-inferiority margins subdivided by medical specialty is shown in Figure 3A. There was a wide range of non-inferiority margins for trial outcomes that involve mortality (0.4 to 30%), with a skewed distribution and distinct peaks observed at 5, 9 and 15%. Thrombosis trials used smaller non-inferiority margins more commonly than did other trials.

Figure 3B illustrates a similarly skewed distribution of relative non-inferiority margins subdivided by medical specialty. The most common relative non-inferiority margin observed was in the range of 1.26 to 1.5. Most relative non-inferiority margins clustered in the range of 1.3 to 1.7, however there were also many relative non-inferiority margins that were greater than 2.

<<**Figure 3: Distribution of absolute and relative non-inferiority margins for primary outcomes involving mortality**>>

Characteristics of non-inferiority margins

The characteristics of the non-inferiority margins in the trials included in this review are summarized in Table 1. The median absolute non-inferiority margin was 9% (IQR 4.2-10%) and the median relative non-inferiority margin was 1.5 (IQR 1.3-1.7).

Table 1: Summary of characteristics of non-inferiority trials included

		Absolute non-inferiority margin (%) for outcomes involving mortality			Relative non-inferiority margin for outcomes involving mortality		
	n (%)	Median (IQR[a])	Range	p-value	Median (IQR[a])	Range	p-value
Overall	111	9 (4.2-10)	0.4-30	--	1.5 (1.3-1.7)	1.1-4.5	--
Medical specialty							
Thrombosis	37 (33.3)	3.6 (2-5)	0.4-30	<0.001	1.4 (1.3-1.7)	1.1-3.9	0.02
Infectious diseases	31 (27.9)	10 (5.5-10)	3-20		1.5 (1.5-1.7)	1.2-4.5	
Oncology	23 (20.7)	10 (7.5-13.8)	4-17.5		1.5 (1.3-1.9)	1.1-3	
Transplant	11 (9.9)	10 (10-13.5)	9-20		1.7 (1.5-2.5)	1.4-4.3	
Cardiology	4 (3.6)	7.5 (5-10.5)	5-12		1.3 (1.2-1.3)	1.1-1.4	
Gastroenterology	2 (1.8)	12.5 (11.3-13.8)	10-15		1.9 (1.6-2.2)	1.3-2.5	
Respirology	2 (1.8)	12.3 (11.1-13.4)	10-14.5		1.3 (1.27-1.3)	1.25-1.32	
Anesthesia	1 (0.9)	10 (10-10)	10-10		1.4 (1.4-1.4)	1.4-1.4	
Pediatric patients included							
Yes	21 (18.9)	10 (5-10)	3.5-15	0.11	1.7 (1.5-2)	1.2-4.5	0.10
No	85 (76.6)	8 (4-10)	0.7-30		1.5 (1.3-1.7)	1.1-4.3	
Unclear/not explicitly stated	5 (4.5)	5.5 (2.2-5.6)	0.4-10		1.4 (1.1-1.5)	1.1-3.2	
Mortality outcome							
Single	23 (20.7)	10 (5-12.8)	0.4-15	0.24	1.3 (1.2-1.6)	1.1-2.5	0.03
Composite	88 (79.3)	7.8 (4-10)	0.8-30		1.5 (1.4-1.7)	1.1-4.5	
Industry funding							
Yes	61 (55)	5.6 (3.5-10)	0.4-30	0.01	1.5 (1.3-1.7)	1.1-4.5	0.41
No	42 (37.8)	10 (5-10)	0.8-15		1.5 (1.4-1.8)	1.1-4	
Unclear/not explicitly stated	8 (7.2)	10 (10-13.4)	5.5-15		1.4 (1.3-2.7)	1.3-4.3	
Pre- and post-2010 release of draft FDA guidance statement							
Pre-2010	35 (31.5)	7.9 (3.8-10)	0.4-15	0.24	1.4 (1.2-1.7)	1.1-4	0.02
Post-2010	76 (68.5)	9.5 (4.5-10)	0.8-30		1.5 (1.4-1.7)	1.2-4.5	

[a]Interquartile range

The differences in both absolute and relative non-inferiority margins used among medical specialties were significant. Thrombosis trials had the lowest median absolute non-inferiority margin of 3.6%. Although there was a wide range of absolute and relative non-inferiority margins used across trials, the absolute non-inferiority margins of at least one trial in every specialty was 10% or greater.

Trials with mortality as part of a composite primary outcome had significantly higher relative non-inferiority margins compared to those with mortality as a single primary outcome. In

contrast, when the non-inferiority margin was expressed as an absolute risk difference, there was no significant difference in the margins between type of mortality outcome. Industry-funded trials had a significantly lower median absolute non-inferiority margin compared to those without industry funding.

In this review, 35 (32%) trials were published before 2010 when the first draft FDA guidance statement about non-inferiority trials was published. The relative non-inferiority margin sizes were significantly larger in trials that were published after 2010. A similar trend was seen with the absolute non-inferiority margin sizes, but the difference was not statistically significant.

Association between trial characteristics and non-inferiority margin size

In our regression analyses of the association between trial characteristics and non-inferiority margin sizes, log-transformation of the non-inferiority margin (outcome variable) resulted in slight improvements to the performance of the regression models. The diagnostic plots of the regression models before and after log-transformation of the absolute and relative non-inferiority margins are provided in Appendix D. Table 2 shows the β coefficients with 95% confidence intervals (CI) from our regression analysis of the association between trial characteristics and the log-transformed absolute non-inferiority margin (adjusted R-squared = 0.44). Thrombosis trials had significantly smaller log-absolute non-inferiority margins compared to trials in infectious diseases (reference group) when adjusted for pediatric patients, single or composite mortality outcome and industry funding. When we analyzed the same for log-transformed relative non-inferiority margins (Table 3), trials related to transplant had

significantly larger log-relative non-inferiority margins compared to infectious diseases (adjusted R-squared = 0.1).

Table 2: Absolute non-inferiority margin regression analyses

Predictor	Adjusted β coefficient*	95% CI	p-value
Specialty			
Oncology	0.12	-0.19 to 0.43	0.45
Cardiovascular	-0.23	-0.84 to 0.39	0.46
Thrombosis	-1.14	-1.48 to -0.8	<0.001
Transplant	0.15	-0.3 to 0.6	0.5
Other[#]	0.42	-0.13 to 0.98	0.14
Infectious diseases	1 (reference)	--	--
Pediatrics			
Yes	-0.22	-0.54 to 0.11	0.19
No	1 (reference)	--	--
Mortality outcome			
Single	-0.27	-0.56 to 0.02	0.07
Composite	1 (reference)	--	--
Industry funding			
Yes	0.08	-0.19 to 0.36	0.55
No	1 (reference)	--	--

*Omnibus F-test: 11.93 (8, 102), p<0.05, adjusted R-squared = 0.44

[#]Due to low number of trials, "Other" category combines trials in anesthesia, gastroenterology and respirology

Table 3: Relative non-inferiority margin regression analyses

Predictor	Adjusted β coefficient*	95% CI	p-value
Specialty			
Oncology	-0.08	-0.24 to 0.07	0.3
Cardiovascular	-0.23	-0.54 to 0.07	0.13
Thrombosis	0.01	-0.16 to 0.18	0.9
Transplant	0.24	0.01 to 0.46	0.04
Other[#]	-0.13	-0.41 to 0.15	0.35
Infectious diseases	1 (reference)	--	--
Pediatrics			
Yes	0.1	-0.06 to 0.26	0.23
No	1 (reference)	--	--
Mortality outcome			
Single	-0.11	-0.25 to 0.03	0.13
Composite	1 (reference)	--	--
Industry funding			
Yes	-0.13	-0.27 to 0.01	0.06
No	1 (reference)	--	--

*Omnibus F-test: 2.56 (8, 102), $p<0.05$, adjusted R-squared = 0.1

[#]Due to low number of trials, "Other" category combines trials in anesthesia, gastroenterology and respirology

DISCUSSION

We conducted a systematic review of 111 non-inferiority trials that compared pharmacological therapies where mortality was included in the primary outcome. We found that the majority of non-inferiority trials focused on thrombosis, infectious diseases and oncology. There was a wide range of non-inferiority margins used in these trials, irrespective of whether they were expressed as a measure of absolute effect or when converted to a relative effect. Our results showed that in the design of at least half of the non-inferiority trials included in this review, all of which included mortality as part of their primary outcome, a "new" drug therapy could have an absolute increase of 9% or relative increase of 50% in mortality outcomes compared to controls and still be accepted as non-inferior. Accepting the potential for this increase in

mortality while declaring non-inferiority is a challenging methodological issue in the conduct of non-inferiority trials.

In our review, we also found that non-inferiority margins were more commonly expressed in terms of absolute risk differences than in relative terms. Whether to present absolute or relative non-inferiority margins is a source of debate in the design of non-inferiority trials.[16] There is no clear consensus on the selection of the most appropriate effect measure but it has been demonstrated that different ways of expressing effect measures could result in different conclusions within the same non-inferiority trial.[14,16-17] Since a relative non-inferiority margin accounts for the estimated baseline risk of outcome, it would be a more conservative choice over an absolute margin to conclude non-inferiority should the event rate in the control group be lower than expected.

We detected significant variations in absolute and relative non-inferiority margin size according to medical specialty which could be partially explained by differences in acuity of diseases, patient age and life expectancy.[18] We also found that industry-funded trials had significantly lower median absolute non-inferiority margins compared to those without industry funding, presumably related to greater financial resources and higher capacity to support larger trials that are necessary when smaller non-inferiority margins are used. However, the difference was not significant when relative non-inferiority margins were compared between trials with and without industry funding.

When we compared non-inferiority margin size between trials published before and after the release of the 2010 draft FDA guidance document on non-inferiority studies, we found that the median non-inferiority margin in trials published after 2010 was increased rather than decreased. This was significant only for relative non-inferiority margins, but not for absolute non-inferiority margins. Perhaps future guidelines could generate reductions in non-inferiority margins used for randomized controlled trials involving mortality, if they recommend margins lower, rather than higher than the current median (<9% absolute mortality, <1.5 times relative mortality).

There is currently limited research in the pre-defined determinants of the size of non-inferiority margins used in non-inferiority trials. Gayet-Ageron et al. conducted a survey among trialists to assess the association of pre-defined trial factors and non-inferiority margins. They found that lower non-inferiority margins were associated with mortality as a primary outcome, low baseline risk and lower costs of new treatments. In contrast, population age group and difficulties with patient recruitment did not appear to affect the choice of margin.[19] Because of the nature of a survey study, these results were based on self-report by respondents and were not necessarily reflective of actual practice when non-inferiority trials are designed and conducted.

In our review of published non-inferiority trials of drug therapies that included mortality as part of their primary outcome, we examined for an association between non-inferiority margin size and medical specialty, inclusion of pediatric patients, mortality as a single or composite

outcome and presence of industry funding. Medical specialty was the only trial characteristic found to be significantly associated with the size of non-inferiority margins--specifically, thrombosis trials were associated with smaller absolute non-inferiority margins, while transplant trials were associated with larger relative non-inferiority margins.

Similar to Gayet-Agergon et al.'s[19] results, we found a significant correlation between the size of absolute non-inferiority margins and estimated baseline risks of outcomes in the control group. While this association was moderate ($r_s = 0.6$), it suggests that larger absolute non-inferiority margins are used when estimated risks of outcome occurring in control groups are higher. As can be seen in Figure 2, this relationship appears most evident for baseline outcome risks up to approximately 20%, beyond which larger absolute non-inferiority margins are no longer associated with higher baseline risk of outcome.

A strength of our review is the comprehensive and sensitive search for non-inferiority trials which spanned a 30-year period to ensure that virtually all non-inferiority trials with a primary outcome involving mortality would be captured. There were no limits placed on the type of medical specialty or patient population as long as a trial compared mortality between pharmacological therapies. However, we relied on authors to provide the values of non-inferiority margins and estimated risks of outcome in their sample size calculations within the publication or in their supplementary materials. The accuracy of reporting these variables was taken at face value. To enable standardized comparisons of absolute and relative non-inferiority margins to be made consistently across all trials included in the review, we omitted

non-inferiority trials that used hazard ratios, odds ratios and event rate measures that either changed with time or would not allow us to determine the estimated risk of outcome in the control group required for analyses. Although there was a large amount of variability in the regression models with low adjusted R-squared values, the direction and significance of the independent variables adjusted for in the models indicated that there was an important effect of medical specialty on non-inferiority margin size.

The absolute and relative non-inferiority margins used in published trials comparing medications are large, allowing conclusions of non-inferiority in the context of large differences in mortality, and highly variable. Most trials utilize non-inferiority margins based on an absolute risk difference, which has only a moderate association with baseline estimates of risk for outcomes. With increasing popularity of non-inferiority trials, clinicians and other users of the medical literature should pay close attention to the size of non-inferiority margins used in these trials and consider the influence of study design parameters and inherent trial characteristics when interpreting the results. A collaborative effort to develop standards for the design and analyses of future non-inferiority trials would be beneficial to the scientific community.

REFERENCES

1 Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006;295:1152-60.

2 Mauri L, D'Agostino RB. Challenges in the design and interpretation of noninferiority trials. *N Engl J Med* 2017;377:1357-67.

3 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.

4 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. doi:10.1371/journal.pone.0013550

5 Schiller P, Burchardi N, Niestroj M, Kieser M. Quality of reporting of clinical non-inferiority and equivalence randomised trials—update and extension. *Trials* 2012;13:214. https://doi.org/10.1186/1745-6215-13-214

6 Lange S, Freitag G. Choice of delta: requirements and reality—results of a systematic review. *Biom J* 2005;47:12-27.

7 Althunian TA, de Boer A, Klungel OH, Insani WN, Groenwold RH. Methods of defining the non-inferiority margin in randomized, double-blind controlled trials: a systematic review. *Trials* 2017;18:107. https://doi.org/10.1186/s13063-017-1859-x

8 Rehal S, Morris TP, Fielding K, Carpenter JR, Phillips PP. Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. *BMJ Open* 2016;6:e012594. doi: 10.1136/bmjopen-2016-012594

9 Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. *Trials* 2011;12:106. https://doi.org/10.1186/1745-6215-12-106

10 D'Agostino RB, Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues—the encounters of academic consultants in statistics. *Statist Med* 2003;22:169-86.

11 Aberegg SK, Hersh AM, Samore MH. Empirical consequences of current recommendations for the design and interpretation of noninferiority trials. *J Gen Intern Med* 2018;33:88-96.

12 Hersh AM, Walter RJ, Abberegg SK. Use of mortality as an endpoint in noninferiority trials may lead to ethically problematic conclusions. *J Gen Intern Med* 2019;34:618-23.

13 Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. Non-inferiority clinical trials to establish effectiveness—guidance for industry. Silver Spring, MD: Food and Drug Administration, November 2016. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/non-inferiority-clinical-trials

14 Head SJ, Kaul S, Bogers AJ, Kappetein AP. Non-inferiority study design: lessons to be learned from cardiovascular trials. *Eur Heart J* 2012;33:1318-24.

15 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.

16 Abulizi X, Flandre P. Choice of treatment-effect measures when noninferiority margins originally defined in absolute difference translated into relative difference influenced the results of clinical trials. *J Clin Epidemiol* 2018;96:63-72.

17 Althunian TA, de Boer A, Groenwold RHH, Klungel OH. Defining the noninferiority margin and analysing noninferiority: an overview. *Br J Clin Pharmacol* 2017;83:1636-42.

18 Gladstone BP, Vach W. Choice of non-inferiority (NI) margins does not protect against degradation of treatment effects on an average—an observational study of registered and published NI trials. *PLoS ONE* 2014;9:e103616. doi:10.1371/journal.pone.0103616

19 Gayet-Ageron A, Agoritsas T, Rudaz S, Courvoisier D, Perneger T. The choice of the noninferiority margin in clinical trials was driven by baseline risk, type of primary outcome, and benefits of new treatment. *J Clin Epidemiol* 2015;68:1144-51.

STATEMENTS

Author contributions

SP, ND and RF drafted the study protocol. SP designed the data collection form, screened titles and abstracts, performed statistical analyses and drafted the study manuscript. SP and MU reviewed full-text articles and performed data collection. ND, MU, RF, NM, WS, JH and MS critically reviewed the final manuscript.

Licence statement

Competing interests declaration

All authors have completed the ICMJE uniform disclosure form and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Ethics approval

Not required

Funding statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. SP is supported by a SickKids Clinician-Scientist Training Program Scholarship from The Hospital for Sick Children. MU is supported by a Vanier Canada Graduate Scholarship from the Canadian Institutes of Health Research.

Patient and public involvement statement

It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Data sharing statement

Data and statistical code are available on request to the corresponding author.

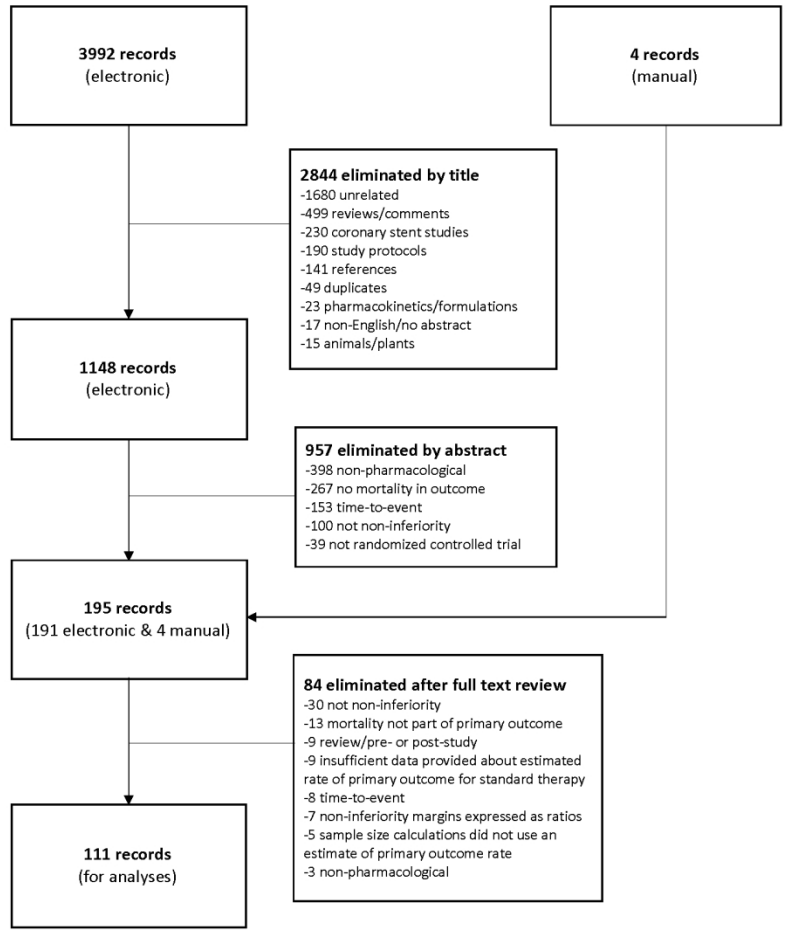


Figure 1: PRISMA flow diagram[14]

Figure 1: PRISMA flow diagram

215x279mm (200 x 200 DPI)

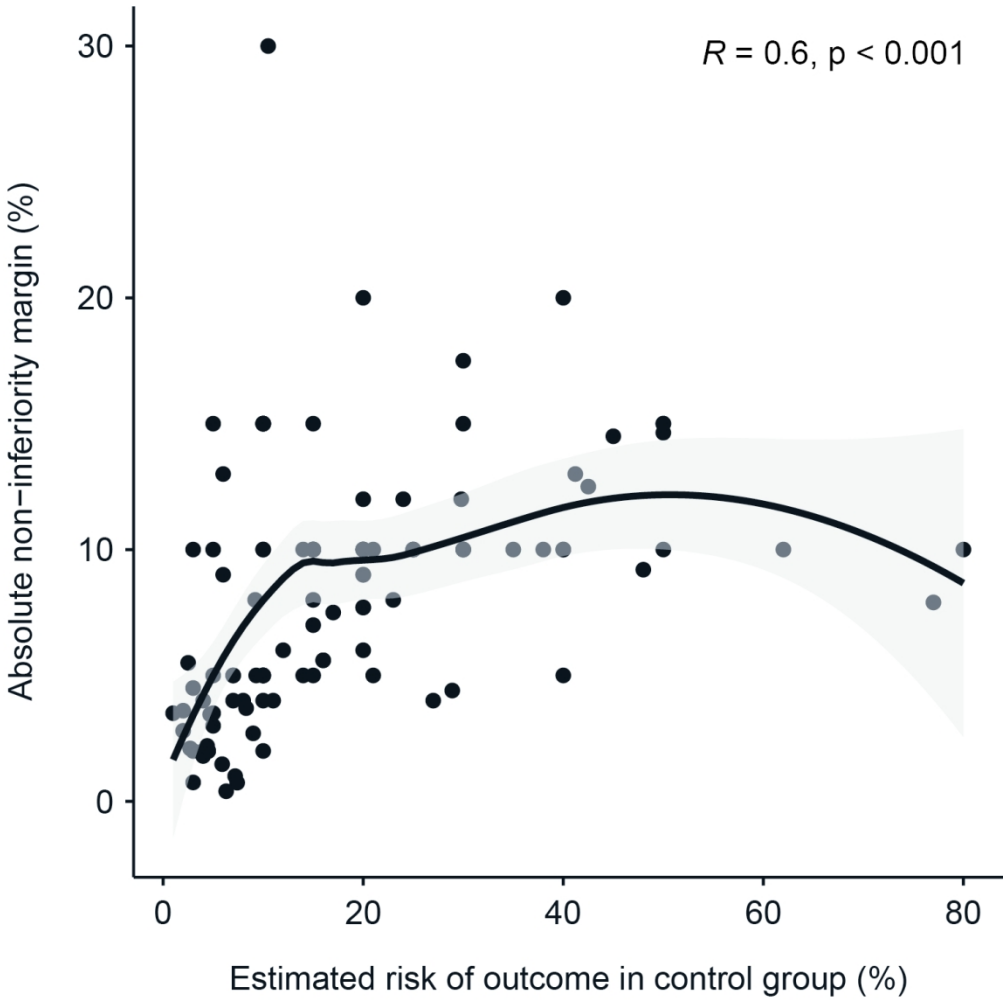


Figure 2: Association between absolute non-inferiority margins and estimated risks of outcome in control group

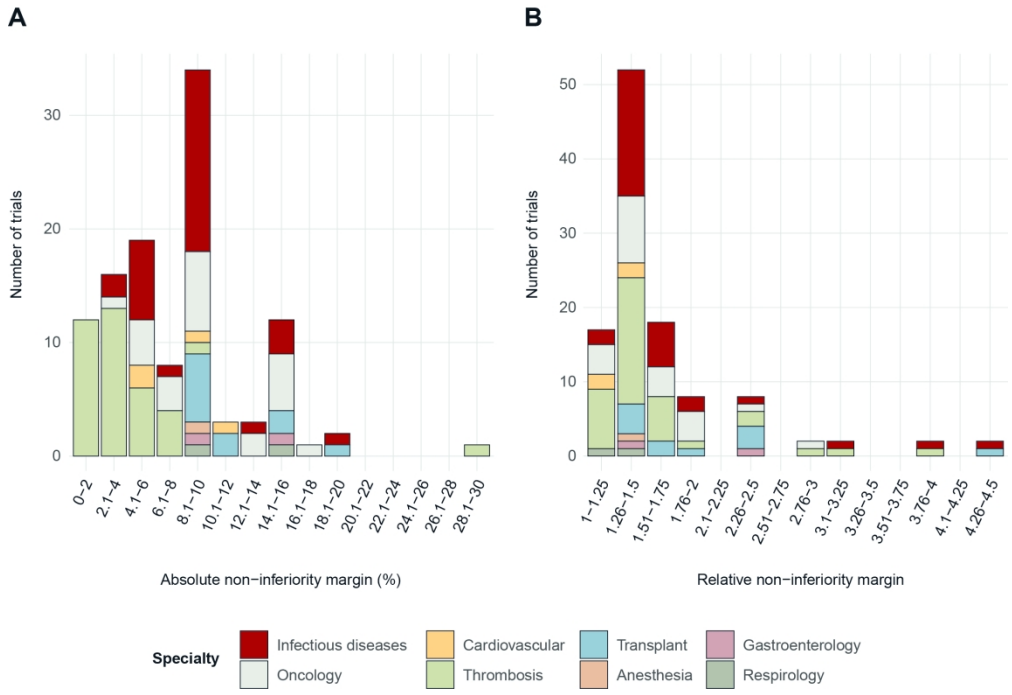


Figure 3: Distribution of absolute and relative non-inferiority margins for primary outcomes involving mortality

Appendix A: Search strategy

MEDLINE

	Searches	Results	Annotations
1	equivalence trial/	503	
2	equivalence trials as topic/	282	
3	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70165	
4	or/1-3	70414	
5	mo.fs.	561893	
6	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493002	
7	mortality/ or cause of death/ or fatal outcome/ or hospital mortality/ or infant mortality/ or maternal mortality/	213608	
8	or/5-7 [****mortality terms****]	1834719	
9	4 and 8 [****base clinical set****]	9115	
10	randomized controlled trial.pt.	500401	
11	controlled clinical trial.pt.	93574	
12	randomized.ab.	469732	
13	placebo.ab.	205181	
14	drug therapy.fs.	2180485	
15	randomly.ab.	327066	
16	trial.ab.	494306	
17	groups.ab.	2008069	
18	or/10-17	4632786	

19	exp animals/ not humans.sh.	4672546
20	18 not 19 [****Cochrane Handbook Highly Sensitive Search Strategy for identifying randomized trials (Box 6.4.c 2008 version)****]	4014138
21	9 and 20 [****Final results****]	3721
22	limit 21 to (english language and humans and yr="1989 -Current")	3212

Medline-in-Process

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	

Medline Epubs Ahead of Print

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	
13	equivalenc*adj3 trial*.ti,ab,kf.	0	
14	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
15	or/13-14	70227	
16	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
17	15 and 16 [****Base clinical set****]	5800	
18	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
19	placebo.ab.	205277	

20	randomly.ab.	327266
21	trial.ab.	494765
22	groups.ab.	2009271
23	or/18-22 [***Trial terms***]	2722197
24	17 and 23 [****Final results****]	2313

Embase Classic+Embase databases (OvidSP)

	Searches	Results	Annotations
1	equivalenc*adj3 trial*.ti,ab,kf.	0	
2	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227	
3	or/1-2	70227	
4	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992	
5	3 and 4 [****Base clinical set****]	5800	
6	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027	
7	placebo.ab.	205277	
8	randomly.ab.	327266	
9	trial.ab.	494765	
10	groups.ab.	2009271	
11	or/6-10 [***Trial terms***]	2722197	
12	5 and 11 [****Final results****]	2313	
13	equivalenc*adj3 trial*.ti,ab,kf.	0	

14	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab,kf.	70227
15	or/13-14	70227
16	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab,kf.	1493992
17	15 and 16 [****Base clinical set****]	5800
18	((randomized or randomised or controlled) adj3 trial*).ti,ab,kf.	401027
19	placebo.ab.	205277
20	randomly.ab.	327266
21	trial.ab.	494765
22	groups.ab.	2009271
23	or/18-22 [***Trial terms***]	2722197
24	17 and 23 [****Final results****]	2313
25	(equivalen* adj2 trial*).ti,ab.	758
26	(noninferiority or non-inferiority or equivalence or equivilency or equivilencies or inferiority or "NI margin*" or "delta margin*" or (prespecified adj2 margin*) or margins).ti,ab.	69786
27	or/25-26 [***equivalency or non-inferiority terms****]	69985
28	(mortality or mortalities or death or deaths or dying or fatal*).ti,ab.	1473665
29	mortality/ or cancer mortality/ or childhood mortality/ or embryo mortality/ or fetus mortality/ or infant mortality/ or maternal mortality/ or prenatal mortality/ or surgical mortality/ or perinatal mortality/ or newborn mortality/	77281
30	death/ or "cause of death"/ or dying/ or heart death/ or sudden death/ or child death/ or newborn death/	75385
31	or/28-30 [****mortality terms****]	1511412

32	27 and 31 [****Base clinical set****]	5823
33	randomized controlled trial/	500622
34	double-blind procedure/	0
35	single-blind procedure/	0
36	crossover-procedure/	0
37	random*.ti,ab,kw.	1109854
38	factorial*.ti,ab,kw.	29313
39	crossover*.ti,ab,kw.	62938
40	"cross over".ti,ab,kw.	22789
41	"cross-over*".ti,ab,kw.	23015
42	placebo*.ti,ab,kw.	212808
43	(doubl* adj5 blind*).ti,ab,kw.	149043
44	(singl* adj5 blind*).ti,ab,kw.	22640
45	assign*.ti,ab,kw.	308920
46	allocat*.ti,ab,kw.	115873
47	volunteer*.ti,ab,kw.	189496
48	or/33-47 [****Cochrane Box 6.3.2.2 EMBASE sensitive TherapyTreatment Effectiveness Filter terms****]	1731635
49	32 and 48 [***Cochrane trial filter****]	1802
50	ct.fs.	0
51	32 and 50 [***clinical trial subheading****]	0
52	limit 32 to (randomized controlled trial or controlled clinical trial or multicenter study)	1230

53	or/49,51-52 [****Final results****]	1974
54	limit 53 to conference abstract [Limit not valid in Ovid MEDLINE(R),Ovid MEDLINE(R) Daily Update,Ovid MEDLINE(R) In-Process,Ovid MEDLINE(R) Publisher; records were retained]	1974
55	53 not 54 [***Conference abstracts removed****]	0
56	limit 55 to (human and english language and yr="1989 -Current")	0

Appendix B: Non-inferiority trials

Author	Year	Country	Specialty	Pediatric patients included	Mortality as single or composite outcome	Industry funding	Conclusion	Estimated risk of event in control (%)	Observed event in control (%)	Absolute non-inferiority margin (% difference)	Relative non-inferiority margin
African Neonatal Sepsis Trial (AFRINEST) group	2015	international	infectious diseases	yes	composite	no	non-inferior	10	8	5	1.5
Ahn et al.	2013	Korea	oncology	no	single	not explicit/unclear	non-inferior	41.2	23	13	1.32
Aliberti et al.	2017	Italy	infectious diseases	no	composite	no	stopped early	10	7.4	5	1.5
Baqui et al.	2015	Bangladesh	infectious diseases	yes	composite	no	non-inferior	10	10	5	1.5
Bertrand et al.	2006	Canada	thrombosis	no	composite	yes	non-inferior	23	18.2	8	1.35
Beyer-Westendorf et al.	2017	Germany	thrombosis	no	composite	yes	non-inferior	3	1.7	4.5	2.5
Borchmann et al.	2017	international	oncology	no	composite	yes	non-inferior	12	8.8	6	1.5
Brack et al.	2012	international	infectious diseases	yes	composite	yes	not non-inferior	1	2.9	3.5	4.5
Budde et al.	2014	international	transplant	no	composite	yes	non-inferior	15	19.6	10	1.67
The Matisse Investigators	2003	international	thrombosis	no	composite	yes	non-inferior	5	5	3.5	1.7
Bunnapradist et al.	2013	international	transplant	no	composite	yes	non-inferior	6	2.5	9	2.5
Cai et al.	2014	China	transplant	no	composite	yes	non-inferior	20	16.7	20	2
Chastre et al.	2003	France	infectious diseases	no	single	no	non-inferior	40	17.2	10	1.25
Cordonnier et al.	2009	France	infectious diseases	no	single	no	non-inferior	9.2	2.7	8	1.87
de Kraker et al.	2004	international	oncology	yes	composite	no	non-inferior	15	8.6	10	1.67
De Simone et al.	2012	international	transplant	no	composite	yes	non-inferior	24	9.5	12	1.5
Diener et al.	2006	international	thrombosis	no	composite	yes	non-inferior	9.3	8.8	5	1.54
Eckardt et al.	2006	international	oncology	no	single	yes	non-inferior	62	69	10	1.16
Eriksson et al.	2011	international	thrombosis	no	composite	yes	non-inferior	20	8.8	7.7	1.39
Eriksson et al.	2007	international	thrombosis	no	composite	yes	non-inferior	48	37.7	9.2	1.19

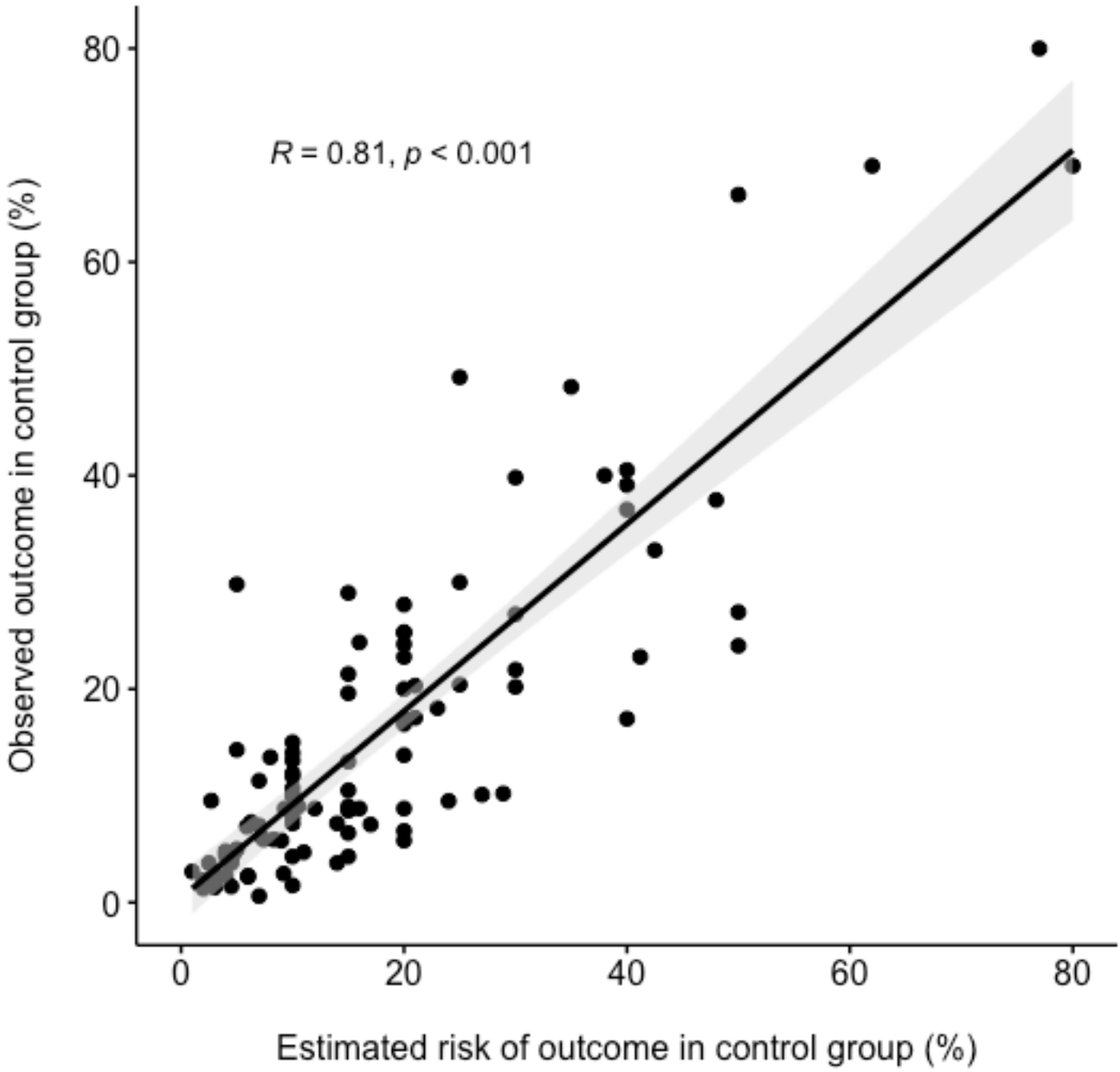
Eriksson et al.	2007	international	thrombosis	no	composite	yes	non-inferior	20	6.7	7.7	1.39
Feres et al.	2013	Brazil	thrombosis	no	composite	yes	non-inferior	9	5.8	2.7	1.3
Ferme et al.	2017	international	oncology	yes	composite	no	non-inferior	10	10.1	10	2
Gaston et al.	2009	US	transplant	yes	composite	yes	non-inferior	20	27.9	10	1.5
Giamarellou et al.	2000	Greece	infectious diseases	no	composite	not explicit/unclear	non-inferior	25	49.2	10	1.4
Gilard et al.	2015	international	thrombosis	no	composite	yes	non-inferior	3	1.5	2	1.67
Gulizia et al.	2008	Italy	cardiology	no	composite	not explicit/unclear	non-inferior	38	40	10	1.26
Gwon et al.	2012	Korea	thrombosis	no	composite	yes	non-inferior	10	4.3	4	1.4
Sibbing et al.	2017	international	thrombosis	no	composite	yes	non-inferior	10.5	9	30	3.86
Han et al.	2016	China	thrombosis	no	composite	yes	non-inferior	8.3	5.9	3.7	1.45
Harris et al.	2018	international	infectious diseases	no	single	no	not non-inferior	14	3.7	5	1.36
Hofheinz et al.	2012	Germany	oncology	no	single	yes	non-inferior	42.5	33	12.5	1.29
Huh et al.	2017	Korea	transplant	no	composite	yes	non-inferior	10	13.3	15	2.5
Iversen et al.	2019	Denmark	infectious diseases	no	composite	no	non-inferior	10	12.1	10	2
Jeng et al.	2018	international	transplant	no	composite	yes	non-inferior	20	5.8	12	1.6
Johnson et al.	2016	international	oncology	no	composite	no	not non-inferior	5	14.3	5	2
Jones et al.	2016	Canada	anesthesia	no	composite	no	non-inferior	25	30	10	1.4
Kim et al.	2012	Korea	thrombosis	no	composite	yes	non-inferior	11	4.7	4	1.36
Park et al.	2013	Korea	thrombosis	no	composite	no	non-inferior	3	1.4	0.75	1.25
Hahn et al.	2018	Korea	thrombosis	no	composite	yes	non-inferior	4.5	4.2	2	1.44
Kirchhof et al.	2012	Germany	cardiology	no	composite	yes	not non-inferior	29.8	na	12	1.4
Kirchhof et al.	2018	international	thrombosis	no	composite	no	non-inferior	17	7.3	7.5	1.44
Kutner et al.	2015	US	cardiology	no	single	no	not non-inferior	21	20.3	5	1.24
Le et al.	2017	Vietnam	infectious diseases	no	single	no	non-inferior	15	6.5	10	1.67
Lee et al.	2018	Korea	thrombosis	no	composite	yes	non-inferior	7	0.6	4	1.57
Lee et al.	2009	UK	oncology	not explicit	single	no	non-inferior	80	69	10	1.13

Liu et al.	2013	China	oncology	yes	single	no	stopped early	15	na	15	2
Maertens et al.	2016	international	infectious diseases	no	single	yes	non-inferior	20	20	10	1.5
Mai et al.	2016	international	oncology	no	composite	yes	non-inferior	50	na	15	1.3
Matsumura-Nakano et al.	2019	Japan	thrombosis	no	composite	yes	not non-inferior	8	13.6	4	1.5
Mavroudis et al.	2016	Greece	oncology	no	composite	no	not non-inferior	15	10.5	7	1.47
Mavroudis et al.	2015	Greece	oncology	no	composite	no	not non-inferior	15	4.3	8	1.53
Merle et al.	2014	international	infectious diseases	no	composite	no	not non-inferior	20	17.2	6	1.3
Mesu et al.	2018	international	infectious diseases	yes	composite	no	non-inferior	6	2.4	13	3.17
Meynard et al.	2018	France	infectious diseases	no	composite	yes	not non-inferior	15	29	10	1.67
Mir et al.	2017	Pakistan	infectious diseases	yes	composite	no	non-inferior	10	11.8	5	1.5
Molloy et al.	2018	international	infectious diseases	no	single	no	non-inferior	15	21.4	10	1.67
Nakamura et al.	2017	Japan	thrombosis	no	composite	no	non-inferior	4.5	1.5	2	1.44
Nathan et al.	2005	Niger	infectious diseases	yes	composite	no	non-inferior	15	9	10	1.67
Park et al.	2014	Korea	gastroenterology	no	single	no	not non-inferior	10	na	15	2.5
Paul et al.	2015	Israel	infectious diseases	no	composite	no	not non-inferior	30	27	15	1.5
Ponticelli et al.	2014	Italy	transplant	no	composite	not explicit/unclear	inconclusive	3	2.8	10	4.33
Postma et al.	2015	Netherlands	infectious diseases	no	single	no	non-inferior	5	na	3	1.6
Pritchard-Jones et al.	2015	international	oncology	yes	composite	no	non-inferior	14	7.4	10	1.71
Pujade-Lauraine et al.	2010	international	oncology	no	composite	yes	non-inferior	77	80	7.9	1.1
Qazi et al.	2017	international	transplant	no	composite	yes	not non-inferior	25	20.4	10	1.4
Reynolds et al.	2010	Uganda	infectious diseases	no	composite	no	non-inferior	5	29.8	15	4
Riess et al.	2010	Germany	thrombosis	no	composite	yes	non-inferior	4.7	4.52	3.45	1.73
Russ et al.	2013	international	transplant	no	composite	yes	unclear	10	14	15	2.5
Schrappe et al.	2018	international	oncology	yes	composite	no	not non-inferior	4	4.4	4	2
Schroder et al.	2004	international	oncology	no	single	no	inconclusive	50	24.04	15	1.3

Schulz-Schupke et al.	2015	international	thrombosis	no	composite	yes	stopped early	10	1.6	2	1.2
Tedesco Silva et al.	2010	international	transplant	no	composite	yes	non-inferior	20	24.2	10	1.5
Sinha et al.	2005	international	respirology	yes	single	yes	non-inferior	45	na	14.5	1.32
Stabile et al.	2008	Italy	thrombosis	not explicit	composite	not explicit/unclear	non-inferior	2.5	3.7	5.5	3.2
Stellbrink et al.	2004	Germany	thrombosis	no	composite	yes	non-inferior	4	4.8	2	1.5
Stets et al.	2019	international	infectious diseases	no	composite	yes	non-inferior	21	17.3	10	1.48
Stone et al.	2007	international	thrombosis	no	composite	yes	not non-inferior	5.9	7.1	1.48	1.25
Swaminathan et al.	2011	India	infectious diseases	no	composite	no	not non-inferior	10	15	15	2.5
The GUSTO V Investigators	2001	international	thrombosis	no	single	yes	non-inferior	7.4	5.9	0.74	1.1
Turpie et al.	2009	international	thrombosis	no	composite	yes	non-inferior	27	10.1	4	1.15
Walsh et al.	2004	international	infectious diseases	yes	composite	yes	non-inferior	50	66.3	10	1.2
Willenheimer et al.	2005	international	cardiology	no	composite	yes	not non-inferior	40	36.8	5	1.13
Yahav et al.	2019	international	infectious diseases	no	composite	not explicit/unclear	non-inferior	35	48.3	10	1.29
Yakoub-Agha et al.	2012	international	oncology	no	single	not explicit/unclear	inconclusive	50	27.2	14.64	1.29
Yang et al.	2018	China	oncology	yes	composite	no	non-inferior	5	na	10	3
Daniels et al.	2019	South Africa	infectious diseases	yes	composite	no	non-inferior	7	11.4	5	1.71
Cisneros et al.	2019	international	infectious diseases	no	single	no	not non-inferior	20	25.3	10	1.5
Hahn et al.	2019	Korea	thrombosis	no	composite	yes	non-inferior	4	2.5	1.8	1.45
Jain et al.	2019	India	respirology	yes	composite	no	not non-inferior	40	39.1	10	1.25
Kollef et al.	2019	international	infectious diseases	no	single	yes	non-inferior	20	25.3	10	1.5
Nunn et al.	2019	international	infectious diseases	no	composite	no	non-inferior	30	20.2	10	1.33
Watanabe et al.	2019	Japan	thrombosis	not explicit	composite	yes	non-inferior	4.4	3.7	2.2	1.5
Charbonnier et al.	1998	international	thrombosis	no	composite	yes	non-inferior	7	7.2	5	1.71
Colombo et al.	2014	international	thrombosis	no	composite	yes	non-inferior	4.5	3.7	2	1.44
Lassen et al.	2010	international	thrombosis	not explicit	composite	yes	non-inferior	16	24.37	5.6	1.35

Lassen et al.	2009	international	thrombosis	no	composite	yes	not non-inferior	16	8.8	5.6	1.35
Le Deley et al.	2014	international	oncology	yes	composite	no	inconclusive	30	21.8	10	1.33
Noguchi et al.	2016	international	oncology	no	composite	yes	non-inferior	30	39.8	17.5	1.58
Patte et al.	1991	international	oncology	yes	composite	not explicit/unclear	non-inferior	10	10.7	15	2.5
Platzbecker et al.	2017	international	oncology	no	composite	no	non-inferior	15	13.2	5	1.33
Raffi et al.	2014	international	infectious diseases	no	composite	yes	non-inferior	20	13.8	9	1.45
Rubinstein et al.	2011	international	infectious diseases	no	composite	yes	non-inferior	40	40.5	20	1.5
Schulman et al.	2009	international	thrombosis	no	composite	yes	non-inferior	2	2.1	3.6	2.8
Schulman et al.	2013	international	thrombosis	no	composite	yes	non-inferior	2	1.3	2.8	2.4
Seo et al.	2014	Korea	gastroenterology	yes	composite	no	non-inferior	30	na	10	1.33
Continuous Infusion versus Double-Bolus Administration of Alteplase (COBALT) Investigators	1997	international	thrombosis	not explicit	single	yes	not non-inferior	6.3	7.53	0.4	1.06
Assessment of the Safety and Efficacy of a New Thrombolytic (ASSENT-2) Investigators	1999	international	thrombosis	no	single	yes	non-inferior	7.2	6.151	1	1.14
Vilas-Boas et al.	2014	Brazil	infectious diseases	yes	composite	no	non-inferior	20	23	9	1.45
International Joint Efficacy Comparison of Thrombolytics (INJECT)	1995	international	thrombosis	no	single	yes	non-inferior	2.7	9.53	2.1	1.78
Nitz et al.	2019	Germany	oncology	no	composite	yes	non-inferior	28.9	10.2	4.4	1.15

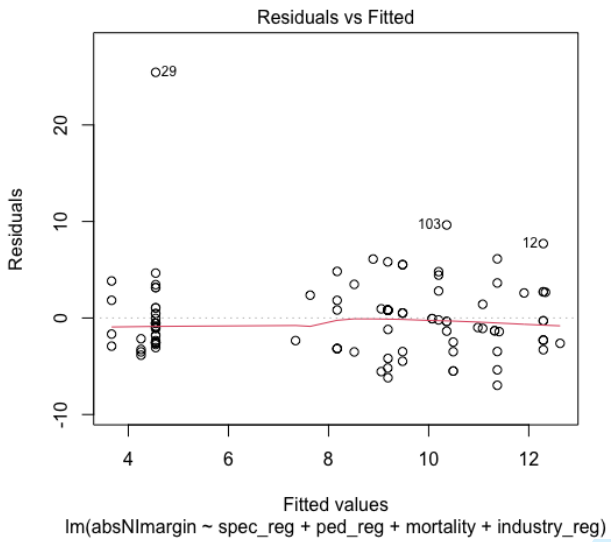
Appendix C: Relationship between observed outcomes and estimated risks of outcome in control group

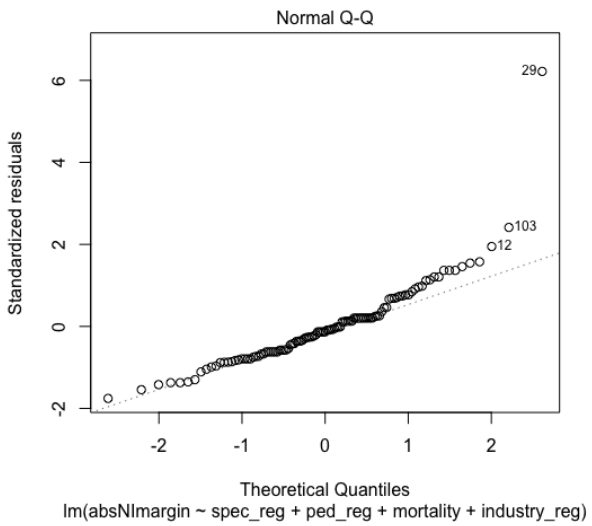


Appendix D: Diagnostic plots for regression models

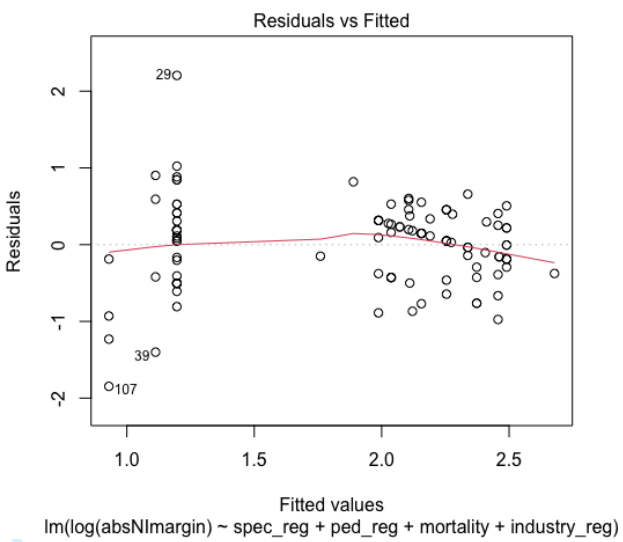
Absolute non-inferiority margins

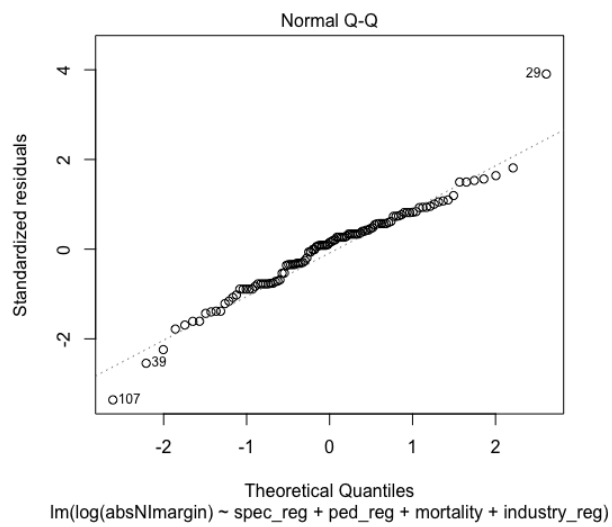
Absolute non-inferiority margin



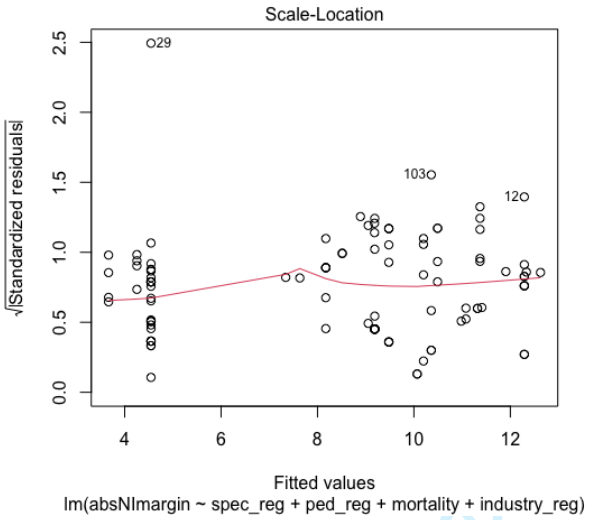


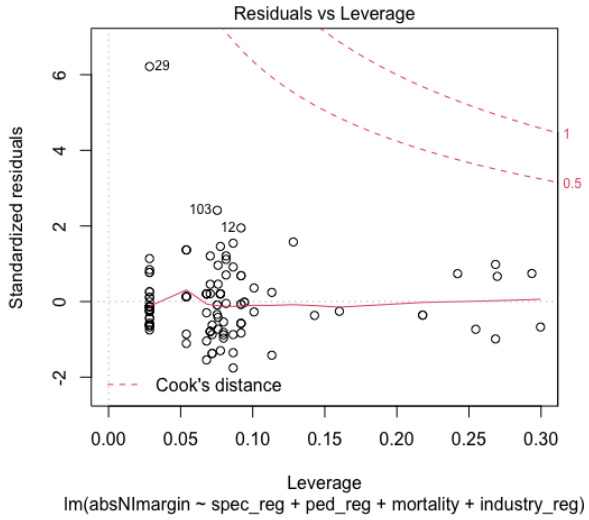
Log-transformed absolute non-inferiority margin



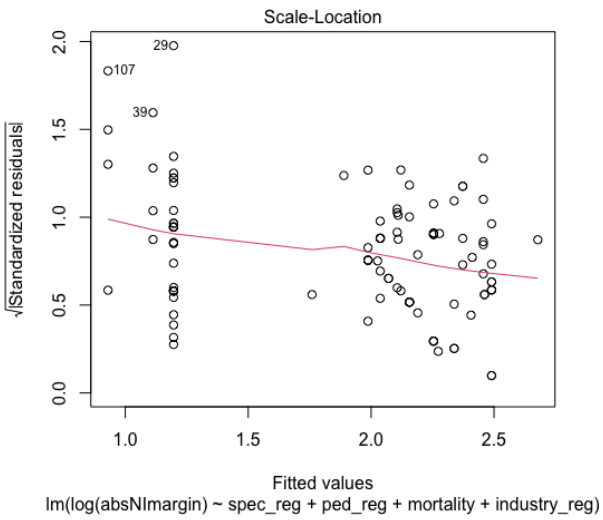


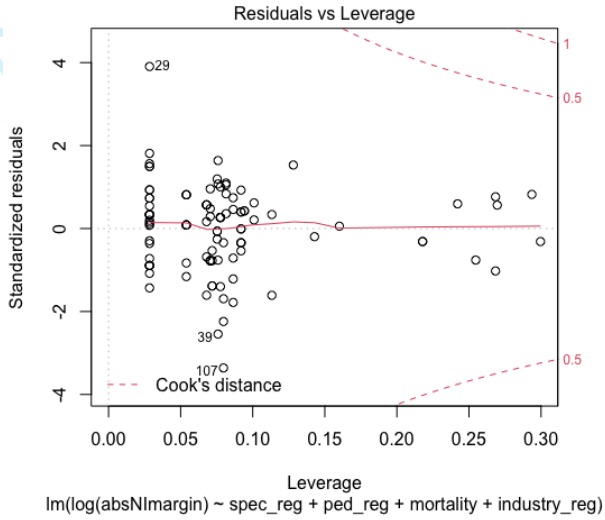
Absolute non-inferiority margin





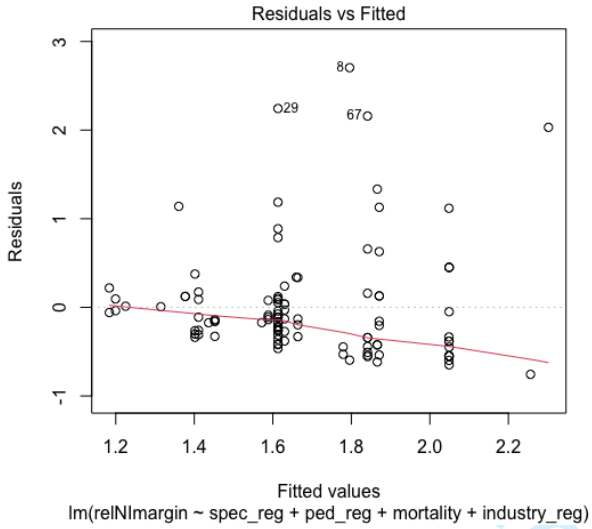
Log-transformed absolute non-inferiority margin

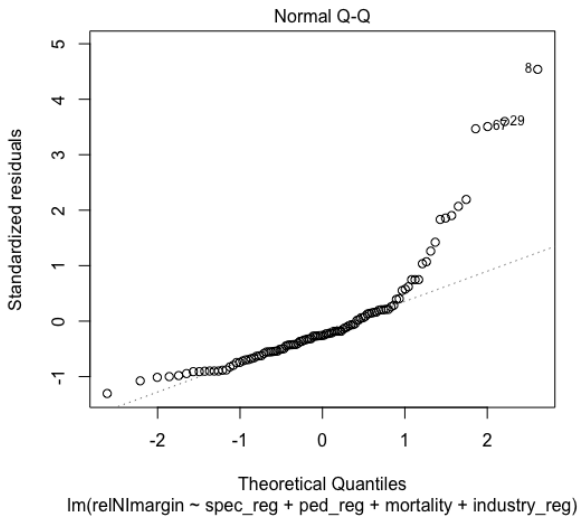




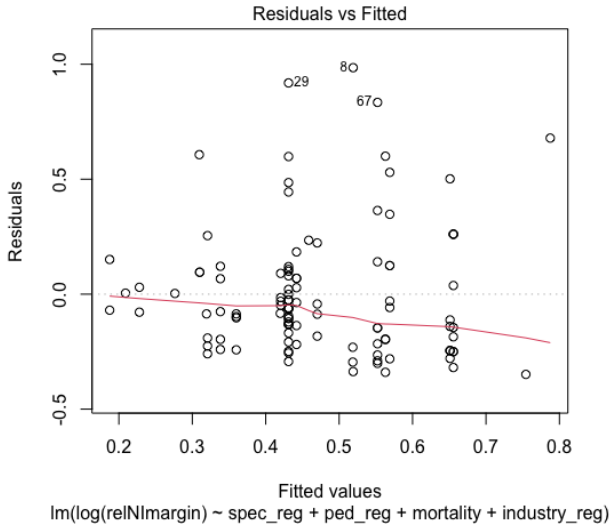
Relative non-inferiority margins

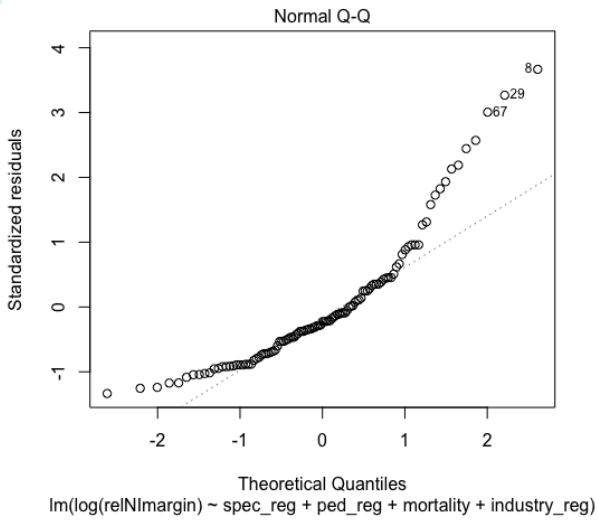
Relative non-inferiority margin



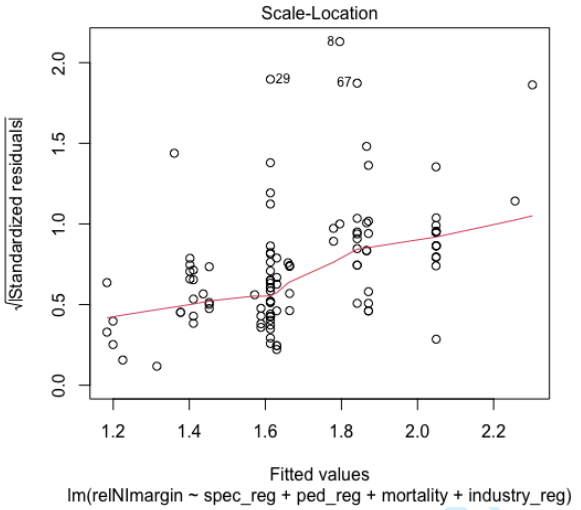


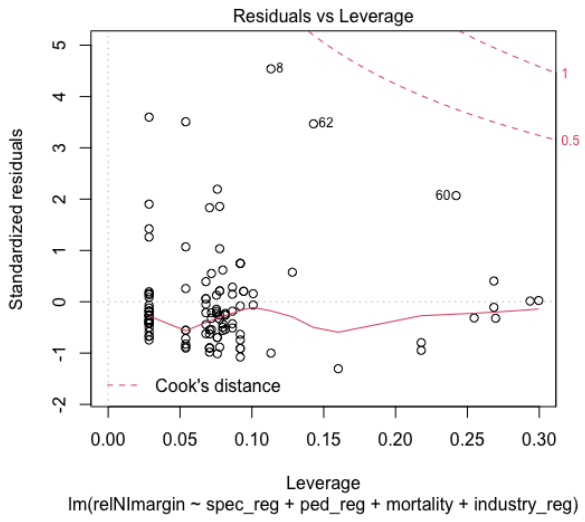
Log-transformed relative non-inferiority margin



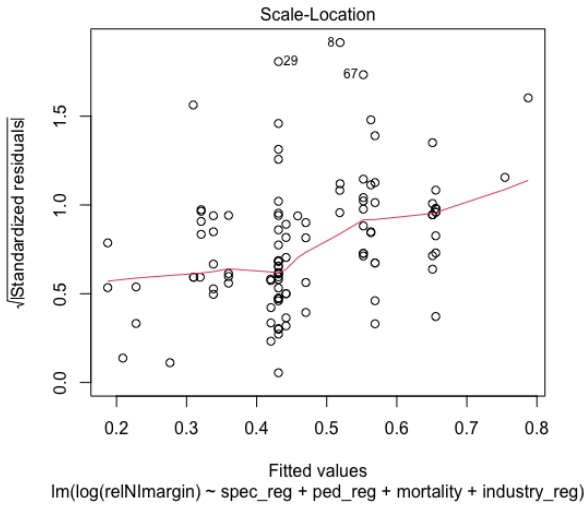


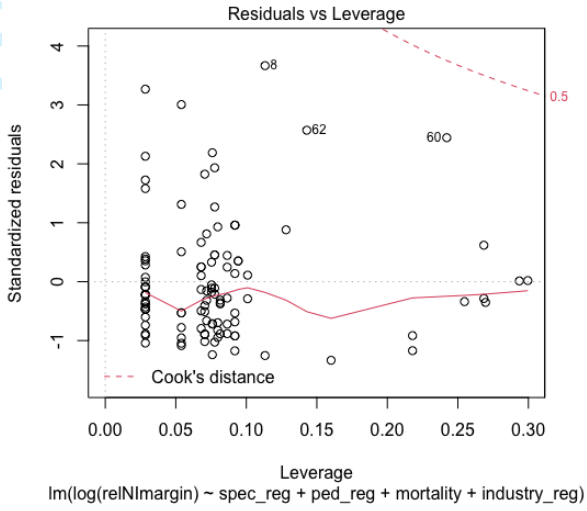
Relative non-inferiority margin





Log-transformed relative non-inferiority margin





PRISMA 2009 Checklist

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

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	n/a
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	n/a
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	10
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	n/a
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	11-15
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	n/a
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	n/a
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	15-18
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	18-19
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	19
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	n/a

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Checklist