

Use of Mortality as an Endpoint in Noninferiority Trials May Lead to Ethically Problematic Conclusions

Andrew M. Hersh, MD¹, Robert J. Walter, MD, DHCE¹, and Scott K. Abberergg, MD, MPH²

¹Division of Pulmonary and Critical Care, Brooke Army Medical Center, Fort Sam Houston, TX, USA; ²Division of Pulmonary and Critical Care Medicine, University of Utah, Salt Lake City, UT, USA.

BACKGROUND: Noninferiority trials are becoming more common. Their design often requires investigators to “trade” a secondary benefit for efficacy. Use of mortality as an outcome of interest leads to important ethical conflicts whereby researchers must establish a minimal clinically important difference for mortality, a process which has the potential to result in problematic conclusions.

OBJECTIVE: We sought to investigate the frequency of the use of mortality as an outcome in noninferiority trials, as well as to determine the average pre-specified noninferiority (“delta”) values.

DESIGN: We searched MEDLINE for reports of parallel-group randomized controlled noninferiority trials published in five high-impact general medical journals.

MAIN OUTCOME MEASURES: Data abstracted from articles including trial design parameters, results, and interpretation of results based on CONSORT recommendations.

RESULTS: One hundred seventy-three manuscripts reporting 196 noninferiority comparisons were included in our analysis. Of these, over a third (67 trials) used mortality either as their sole endpoint (11 trials) or as part of a composite endpoint (56 trials). Nine trials were consort A, 21 trials consort B, 19 trials consort C, 12 were consort F, 4 consort G, and 2 were consort H. Four analyses showed statistically significant more deaths in the new treatment arm, while meeting consort criteria as “inconclusive” (consort G), (Behringer et al. in *Lancet*. 385(9976):1418–1427, 2015; Kaul et al. in *N Engl J Med*. 373(18):1709–1719, 2015; Bwakura-Dangarembizi et al. in *N Engl J Med*. 370(1):41–53, 2014) and thirteen trials utilizing mortality as an endpoint and had an absolute increase of > 3%, and six had an absolute increase of > 5%.

CONCLUSIONS: The use of mortality as an outcome in noninferiority trials is not rare and scenarios where the new treatment is statistically worse, but a conclusion of noninferiority or inconclusive do occur. We highlight these issues and propose simple steps to reduce the risk of ethically dubious conclusions.

KEY WORDS: noninferiority trials; medical ethics; clinical trials; outcomes measures.

J Gen Intern Med 34(4):618–23

DOI: 10.1007/s11606-018-4813-z

© Society of General Internal Medicine 2019

Received July 16, 2018

Revised October 19, 2018

Accepted December 13, 2018

Published online February 12, 2019

INTRODUCTION

Noninferiority trials are increasing in frequency.^{1–3} These trials compare new therapeutic or diagnostic strategies to the ones of established efficacy. This form of trial was first developed as a tool for comparing interventions when one has an intrinsic advantage.^{4, 5} If, for example, an intervention is cheaper, less toxic, more durable, or easier to use, it may not need to be clearly more efficacious—only not worse. In such cases, it may be reasonable to utilize a less stringent standard when performing our statistical testing.

Figure 1 and Table 1 summarize the CONSORT recommendations regarding the establishment of noninferiority, superiority, inferiority, and inconclusive results.⁶ Under this rubric, a noninferiority trial can result in eight possible designations (see Table 1).

Superiority and inferiority are established in the typical manner with an upper or lower confidence interval excluding zero. Noninferiority, however, is established when the confidence intervals exclude a prespecified margin, sometimes called the prespecified margin of indifference, noninferiority margin, or simply delta (see Fig. 1). This margin is selected so that any differences between the new and old interventions are small enough to be considered clinically unimportant. While it is well known that findings may be statistically, but not clinically, significant, noninferiority trials add rigor by requiring pre-specified clinically meaningful differences. An understanding of what constitutes clinically meaningful is therefore integral to the process of deciding upon what noninferiority margins should be tolerated.

Pre-established noninferiority margins are often reported without explanation or with incomplete justification.⁷ Furthermore, many trials include death as an outcome, either primarily or as part of a composite endpoint. Such an arrangement has the potential to result in some disquieting conclusions. An intervention may be found to result in statistically more deaths, and yet still be viewed as “inconclusive” (see Fig. 1, scenario G), or even “noninferior” (see Fig. 1, scenario D). An important question is what constitutes a minimal clinically significant difference when death is the outcome. Unlike other outcomes, the selection of clinically insignificant differences for mortality may be difficult to establish. This study’s purpose was to investigate the frequency of mortality as an outcome in noninferiority trials, determine what previous researchers have

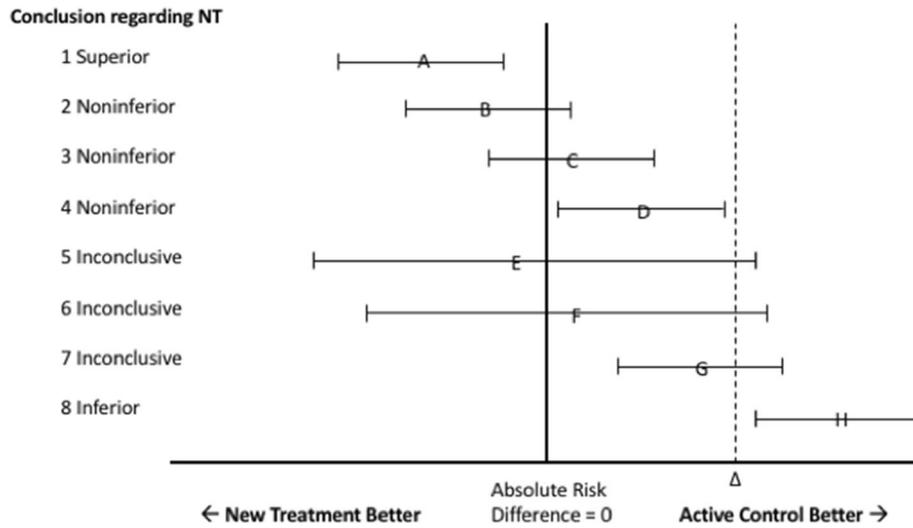


Fig. 1 CONSORT diagram for interpreting the results of noninferiority trials based on the 95% confidence interval of the difference between the new treatment and active control. NT new treatment. The triangle represents the prespecified margin of noninferiority.

specified as clinically insignificant, and report the range of findings in an important subset of noninferiority trials.

METHODS

Our search strategy and data extraction method and validation have been previously described in detail elsewhere.⁷ Briefly, we searched MEDLINE for reports of parallel-group randomized controlled noninferiority trials published in five high-impact general medical journals. We reviewed the resulting titles and abstracts to identify articles that met our principal inclusion criteria: prospective, parallel group, randomized controlled trials where a primary outcome of mortality (either alone or as part of a composite endpoint) was tested using a noninferiority hypothesis. Composite outcomes which required a pre-condition survival (such as progression-free survival) were also included. We reviewed full-text manuscripts for studies passing these initial criteria and excluded trials that had a cluster randomized design, trials where the data were incomplete or could not be summarized, and those that used Bayesian methodology.

We abstracted journal, year of publication, specialty orientation of the trial, treatment in each arm, and primary outcome and whether the primary outcome or part thereof included

mortality. We abstracted methodological data including pre-specified noninferiority margin and its justification or lack. A second author performed an independent data abstraction on a 10% random sample to ensure accuracy and found no discrepancies.

We characterized trial results in terms of the point estimate and confidence intervals in accord with the CONSORT statement and the authors’ conclusions regarding declarations of noninferiority, superiority, inferiority, and inconclusive results (see Fig. 1). If the one-sided alpha used for the analysis was greater than 0.025, we recalculated, where possible, a 2-sided 95% confidence interval for comparison to CONSORT recommendations and documented whether the conclusions were different than the original alpha. For trials that reported co-primary or multiple outcomes, we used the first mentioned outcome as the primary outcome. Some trials compared multiple treatments, e.g., multiple doses of a new drug against one comparator group and we considered these to represent separate comparisons. In determining if justification for the selection of a noninferiority margin value was presented, we coded trials as having “none” if no mention was made as to how it was selected, “abstract” if some mention was made but it was vague or irreproducible, and “concrete” if an explicit reproducible justification was provided.

Table 1 Consort Classification and Description Thereof as well as Number of Trials in Our Cohort in Each Class. Significance Defined as $p < 0.05$; NT new treatment, AC active control

Consort classification	Consort designation	Consort description	Trials, n
A	Superior	Point estimate significantly favors NT, and delta is excluded	7
B	Noninferior	Point estimate nonsignificantly favors NT, delta is excluded	15
C	Noninferior	Point estimate nonsignificantly favors AC, delta is excluded	15
D	Noninferior	Point estimate significantly favors AC, but delta is excluded:	0
E	Inconclusive	Point estimate nonsignificantly favors NT, delta is not excluded	0
F	Inconclusive	Point estimate nonsignificantly favors AC, delta is not excluded:	12
G	Inconclusive	Point estimate significantly favors AC, delta is not excluded:	4
H	Inferior	Point estimate significantly favors AC, delta is excluded by lower bound CI	2

We pooled the reported mortality using a random effects approach. When pre-specified deltas were given as a relative measure, such as hazard ratio, we converted them to absolute risk reduction for use in direct comparison (Fig. 2). A weighted average for the noninferiority margin was calculated. For other relevant clinical variables, we compared central tendency and frequency using the two-sample *t* test or Mann-Whitney *U* test where appropriate for continuous variables, and the χ^2 test or the Fisher's exact test when sample size was small for dichotomous covariates. All displayed *p* values are two-sided with a significance level of $p < 0.05$ (STATA v. 15.1, College Station, TX).

RESULTS

One hundred seventy-three manuscripts reporting 196 noninferiority comparisons met inclusion criteria. Of these, over a

third (67 trials) used mortality either as their sole endpoint (11 trials) or as part of a composite endpoint (56 trials). Trials utilizing mortality as an endpoint were most frequently found in cardiology, oncology, and infectious disease trials. Approximately two thirds of these were industry-sponsored trials, which were more likely to utilize mortality as an endpoint ($p = 0.03$).

The mean pooled pre-specified noninferiority margin for these trials ranged from 0.4 to 19.1% with a mean of 2.8%. This noninferiority margin was more conservative than trials that did not utilize mortality as an endpoint that had an average delta 10.0% ranging from 0.6–25% ($p < 0.001$). Of the 67 trials, only 15 had concrete and explicit justification for their choice of delta, and an additional 3 had incomplete justification, with the remaining 49 trials offering no justification for their chosen noninferiority margin.

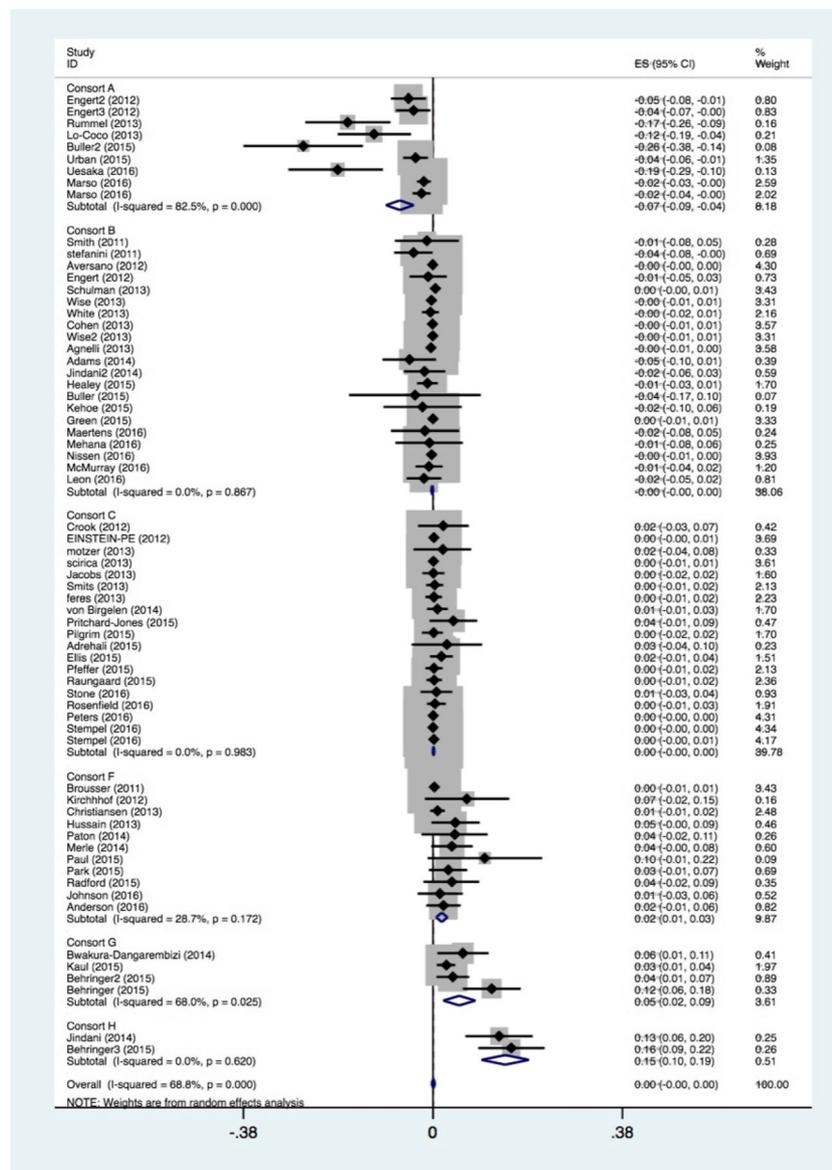


Fig. 2 Forest plot for mortality outcomes by consort designation. All outcomes are reported as absolute risk reduction (ARR).

Table 1 describes trials with mortality as an endpoint as well as their consort classification. We found 9 trials were consort A (point estimate significantly favors NT, and delta is excluded), 21 trials were consort B (point estimate non-significantly favors NT, delta is excluded), 19 trials were consort C (point estimate non-significantly favors old, but “noninferior”), 12 were consort F (point estimate non-significantly favors old, but “inconclusive”), 4 were consort G (point estimate significantly favors old, but “inconclusive”), and 2 were consort H (point estimate significantly favors old).

Mean absolute risk reduction for trials utilizing mortality as an endpoint was 0.1% [95% CI (−0.3%, 0.4%)], where negative risk reduction favors the new intervention, and was similar to trials not utilizing mortality as an endpoint ($p = 0.96$) (Table 2). In our dataset, 12 studies had a non-significantly higher mortality, with an “inconclusive” noninferiority conclusion (consort F).^{23–33, 38} Additionally, four analyses showed statistically significantly greater mortality in the new treatment arm, while meeting consort criteria as “inconclusive” (consort G),^{34–36} and 13 trials utilizing mortality as an endpoint (either alone or in composite) had an

absolute increase of >3%, while 6 had an absolute increase of >5%.

DISCUSSION

Noninferiority trials are important but must be undertaken with a clear understanding of their strengths and limitations. We previously demonstrated that the structure of noninferiority trials could create bias favoring the new treatment.^{7, 39} Trials that utilize mortality as an outcome have an additional ethical obligation to carefully define what constitutes a “clinically insignificant” difference. Otherwise, they may result in conclusions which are counter to ethical reasoning (increased mortality but noninferiority). This is because the logic of noninferiority requires us to establish a threshold (noninferiority margin) below which any difference would be “clinically insignificant.” Such a threshold is difficult to establish when mortality is part of the outcome. If, in fact, one of the proper ends of medicine is fostering the health and well-being of each individual patient, then a lack of meaningful acknowledgement of *any* difference

Table 2 Trials Utilizing Mortality as an Endpoint Whose Point Estimate Favor the Active Control. Negative Risk Reduction Favors the New Intervention. Delta is Reported as Absolute Risk Reduction

Author	Endpoint	Delta (%)	ARR (95% CI)	Consort class
Schulman ⁸	Recurrent or fatal VTE	2.8	0.5 (−0.3 to 1.5)	C
Motzer ⁹	Progression-free survival	14.5	2.0 (−3.9 to 7.7)	C
von Birgelen ¹⁰	Target vessel failure (includes death)	3.6	0.9 (−1.2 to 3.0)	C
Pilgrim ¹¹	Death, MI, revascularization	3.5	0.14 (−2.2 to 2.0)	C
Büller ¹²	Symptomatic recurrent VTE (includes death)	1.8	0.2 (−0.5 to 1.0)	C
Raungaard ¹³	Death, MI, or revascularization	2.5	0.25 (−1.3 to 1.8)	C
Feres ¹⁴	NACCE (includes death, myocardial infarction, stroke, or major bleeding)	2.7	0.2 (−1.5 to 1.8)	C
Jacobs ¹⁵	MACE (includes death, myocardial infarction, revascularization, and stroke)	4.7	0.09 (−2.1 to 2.3)	C
Ellis ¹⁶	Cardiac death, target-vessel myocardial infarction, or ischemia-driven target-lesion revascularization	4.5	1.7 (−0.6 to 4.0)	C
Smits ¹⁷	Death, MI, revascularization	4.0	0.36 (−1.4 to 2.1)	C
Pritchard-Jones ¹⁸	Event-free survival at 2 years after diagnosis	10.0	4.1 (−0.7 to 8.9)	C
Rosenfield ¹⁹	Death, stroke, or MI	3.0	0.4 (−1.2 to 2.6)	C
Crook ²⁰	Survival	8.0	2.1 (−3.0 to 7.2)	C
Ardreghali ²¹	30-day patient/graft survival	10.0	2.8 (−4.3 to 9.9)	C
Stone ²²	Death, stroke, MI at 3 years	4.0	0.7 (−2.5 to 3.9)	C
Merle ²³	Treatment failure, recurrence, or death or study dropout during treatment	6.0	3.8 (−0.4 to 8.0)	F
Christiansen ²⁴	Death, MI, stent thrombosis, revascularization	2.0	0.92 (−0.6 to 2.3)	F
Paton ²⁵	“Good HIV control” (see text, included survival)	10.0	4.4 (−2.3 to 11.0)	F
Kirchhhof ²⁶	Time to persistent atrial fibrillation or death	12.0	6.8 (−1.6 to 15.3)	F
Johnson ²⁷	3-year progression free survival	5.0	1.4 (−3.1 to 6.0)	F
Radford ²⁸	Progression-free survival	7.0	3.8 (−1.9 to 9.3)	F
Park ²⁹	Death, MI, revascularization	4.0	3.1 (−0.8 to 6.9)	F
Anderson ³⁰	Death or disability at 90 days	3.3	2.1 (−1.3 to 5.5)	F
Bousser ³¹	Fatal or non-fatal ischemic stroke, fatal or non-fatal myocardial infarction, or other vascular death	0.55	0.29 (−0.6 to 1.1)	F
Paul ³²	MRSA infections treatment failure at day 7 including death	15	10.4 (−1.1 to 21.9)	F
Hussain ³³	Survival	11.6	4.5 (−0.3 to 9.4)	F
Behringer ³⁴ (analysis #1)	Freedom from treatment failure	6.0	11.8 (6.0 to 17.6)	G
Behringer ³⁴ (analysis #2)	Freedom from treatment failure	6.0	4.0 (0.7 to 7.3)	G
Kaul ³⁵	Cardiac death, target-vessel myocardial infarction, or ischemia-driven target-vessel revascularization	4.0	2.7 (0.78 to 4.5)	G
Bwakura-Dangarembizi ³⁶	Hospitalization or death	3.0	6.0 (0.9 to 11.2)	G
Behringer ³⁴ (Analysis #3)	Freedom from treatment failure	6.0	15.7 (9.1 to 22.4)	H
Jindani ³⁷	“Unfavorable response” (see text)	6.0	13.3 (6.5 to 20.0)	H

HIV human immunodeficiency virus, MI myocardial infarction, MRSA methicillin-resistant *Staphylococcus aureus*, NACCE net adverse cardiac and cerebral events, MACE major adverse cardiac event, VTE venous thromboembolism

in mortality, even within the context of noninferiority, is ethically dubious and such findings should be a central consideration in the discussion of these studies and how their results may impact clinical practice.

Gladstone and Vach previously have described a method they term the advantage deficit assessment (ADA)⁴⁰ where the loss of efficacy is explicitly compared to the secondary advantage gained. Utilizing such a method (even conceptually), one is forced to clarify what secondary advantage (ease of use, reduced cost, etc.) would be necessary to justify a loss in efficacy (in this case higher mortality). Such a cost-benefit analysis is not new to medical decision-making, either in research or clinical practice, and attempts have been made previously to quantify such reasoning. For example, quality-adjusted-life-years (QALYs) are one common conceptualization used to compare the cost-effectiveness of treatments (cost/QALY). While this form of analysis allows for ease of comparison, it raises a litany of ethical concerns which have been voiced elsewhere.⁴¹ An advantage deficit analysis can be thought of in similar terms, where a secondary benefit is considered “per life-year.” Formulating this kind of analysis explicitly is extremely difficult and ethically fraught, and we do not suggest it necessary to attempt to precisely calculate such a metric. Nevertheless, it seems worthwhile, both for researchers undertaking and clinician evaluating a study, to consider what benefit might be required to justify an increase in mortality. A proposed benefit may in fact be sufficiently trivial that no increase in mortality may be justified. In such a case, a noninferiority trial would be inappropriate.

Our findings demonstrate that the use of mortality as an outcome in noninferiority trials is not rare (approximately one third of our analyzed cohort), and scenarios where the new treatment is statistically worse, but a conclusion of noninferiority or inconclusivity do occur—i.e., consort G or H (6/67). One simple solution to this dilemma would be to establish a priori the accepted noninferiority margin with a carefully reasoned explanation of why this particular value was chosen, given the nature of the outcome (death). One would then perform a two-step process whereby trials would be judged first according to standard statistical reasoning (i.e., via a two-sided alpha of < 0.05), and then, only if superiority or inferiority is not established, formal noninferiority can be tested. This approach appears to be what many authors undertake de facto. Authors of the six consort G and H analyses uniformly concluded the new treatment to be inferior. Such an approach would not necessarily be exclusive to trials which utilize mortality as an endpoint, and evidence exists that performing both tests would not incur the “statistical penalty” normally associated with multiple hypothesis testing.⁵

Our study has several limitations. We include only a limited sample of noninferiority trials, albeit the ones from an important set of journals (the five highest impact general medicine publications). Nevertheless, the possibility of publication bias exists given our limited sample and narrow journal selection,

and our sample may not reflect the overall rate in the use of mortality as an endpoint in noninferiority trials. Future studies should look to expand upon our analysis using a larger, more inclusive, sample.

In addition, 11 trials in our cohort used mortality as a sole primary endpoint with a prespecified delta as high as 11.6%;³³ the remainder of trials in our cohort included mortality as part of a composite. In such cases, the contribution of death to the overall outcome is often unclear and may in point-of-fact be small. Such a scenario further complicates the ability to draw useful conclusions as an increase in an unfavorable composite endpoint may be driven largely but a “less important” component (for example coronary revascularization in the case of major adverse cardiac event). Such complexity requires that readers consider each trial on case-by-case basis, but it may be helpful to ask what increase in mortality might a researcher or reader deem acceptable for the proposed benefit. Despite these limitations, we feel our analysis adds a novel perspective on an important and expanding group of trials.

CONCLUSION

Noninferiority trials are an important methodology in our armamentarium of clinical investigations. They must however be undertaken with full knowledge of their limitations—chief among them is the fact that they ask the investigator to “trade” some secondary benefit for efficacy. Such a trade may be justified when the loss of efficacy can reasonably be viewed as clinically insignificant. This however does not appear to be the case when mortality is an endpoint of interest. Our data show that a mortality endpoint is often used in noninferiority trials, and that the seemingly contradictory scenario where a new treatment is simultaneously inferior to (in terms of a superiority test), but inconclusive, does occur. Given the difficulty in clearly defining a minimum clinically significant threshold for mortality—the use of mortality as an endpoint must be undertaken in noninferiority trials with caution. At the very least, studies that include mortality must be careful in selecting and defending what is an acceptable “clinically insignificant” difference.

Corresponding Author: Andrew M. Hersh, MD; Division of Pulmonary and Critical Care Brooke Army Medical Center, Fort Sam Houston, TX, USA (e-mail: Andrew.m.hersh.mil@mail.mil).

Author Contribution Contributions made to the following manuscript by the listed authors include conception and design: AH, RW, and SA; analysis and interpretation of data: AH and SA; and drafting and editing of the manuscript: AH, RW, and SA.

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they do not have a conflict of interest.

Publisher's Note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. **Soonawala D, Middelburg RA, Egger M, Vandenbroucke JP, Dekkers OM.** Efficacy of experimental treatments compared with standard treatments in non-inferiority trials: a meta-analysis of randomized controlled trials. *Int J Epidemiol.* 2010;39(6):1567–1581.
2. **Le Henaff A, Giraudeau B, Baron G, Ravaud P.** Quality of reporting of noninferiority and equivalence randomized trials. *JAMA.* 2006;295(10):1147–1151.
3. **Flacco ME, Manzoli L, Boccia S, et al.** Head-to-head randomized trials are mostly industry sponsored and almost always favor the industry sponsor. *J Clin Epidemiol.* 2015;68(7):811–820.
4. **Garattini S, Bertele V.** Non-inferiority trials are unethical because they disregard patients' interests. *Lancet.* 2007;370(9602):1875–1877.
5. **Lesaffre E.** Superiority, equivalence, and non-inferiority trials. *Bull NYU Hosp Jt Dis.* 2008;66(2):150–154.
6. **Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C.** Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA.* 2012;308(24):2594–2604.
7. **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(1):88–96.
8. **Schulman S, Kearon C, Kakkar AK, et al.** Extended use of dabigatran, warfarin, or placebo in venous thromboembolism. *N Engl J Med.* 2013;368(8):709–718.
9. **Motzer RJ, Hutson TE, Cella D, et al.** Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N Engl J Med.* 2013;369(8):722–731.
10. **von Birgelen C, Sen H, Lam MK, et al.** Third-generation zotarolimus-eluting and everolimus-eluting stents in all-comer patients requiring a percutaneous coronary intervention (DUTCH PEERS): a randomised, single-blind, multicentre, non-inferiority trial. *Lancet.* 2014;383(9915):413–423.
11. **Pilgrim T, Heg D, Roffi M, et al.** Ultrathin strut biodegradable polymer sirolimus-eluting stent versus durable polymer everolimus-eluting stent for percutaneous coronary revascularisation (BIOSCIENCE): a randomised, single-blind, non-inferiority trial. *Lancet.* 2014;384(9960):2111–2122.
12. **Büller HR, Prins MH, Lensin AW, et al.** Oral rivaroxaban for the treatment of symptomatic pulmonary embolism. *N Engl J Med.* 2012;366(14):1287–1297.
13. **Raungaard B, Jensen LO, Tilsted HH, et al.** Zotarolimus-eluting durable-polymer-coated stent versus a biolimus-eluting biodegradable-polymer-coated stent in unselected patients undergoing percutaneous coronary intervention (SORT OUT VI): a randomised non-inferiority trial. *Lancet.* 2015;385(9977):1527–1535.
14. **Feres F, Costa RA, Abizaid A, et al.** Three vs twelve months of dual antiplatelet therapy after zotarolimus-eluting stents: the OPTIMIZE randomized trial. *JAMA.* 2013;310(23):2510–2522.
15. **Jacobs AK, Normand SL, Massaro JM, et al.** Nonemergency PCI at hospitals with or without on-site cardiac surgery. *N Engl J Med.* 2013;368(16):1498–1508.
16. **Ellis SG, Kereiakes DJ, Metzger DC, et al.** Everolimus-Eluting Bioresorbable Scaffolds for Coronary Artery Disease. *N Engl J Med.* 2015;373(20):1905–1915.
17. **Smits PC, Hofma S, Togni M, et al.** Abluminal biodegradable polymer biolimus-eluting stent versus durable polymer everolimus-eluting stent (COMPARE II): a randomised, controlled, non-inferiority trial. *Lancet.* 2013;381(9867):651–660.
18. **Pritchard-Jones K, Bergeron C, de Camargo B, et al.** Omission of doxorubicin from the treatment of stage II-III, intermediate-risk Wilms' tumour (SIOP WT 2001): an open-label, non-inferiority, randomised controlled trial. *Lancet.* 2015;386(9999):1156–1164.
19. **Rosenfield K, Matsumura JS, Chaturvedi S, et al.** Randomized Trial of Stent versus Surgery for Asymptomatic Carotid Stenosis. *N Engl J Med.* 2016;374(11):1011–1020.
20. **Crook JM, O'Callaghan CJ, Duncan G, et al.** Intermittent androgen suppression for rising PSA level after radiotherapy. *N Engl J Med.* 2012;367(10):895–903.
21. **Ardehali A, Esmailian F, Deng M, et al.** Ex-vivo perfusion of donor hearts for human heart transplantation (PROCEED II): a prospective, open-label, multicentre, randomised non-inferiority trial. *Lancet.* 2015;385(9987):2577–2584.
22. **Stone GW, Sabik JF, Serruys PW, et al.** Everolimus-Eluting Stents or Bypass Surgery for Left Main Coronary Artery Disease. *N Engl J Med.* 2016;375(23):2223–2235.
23. **Merle CS, Fielding K, Sow OB, et al.** A four-month gatifloxacin-containing regimen for treating tuberculosis. *N Engl J Med.* 2014;371(17):1588–1598.
24. **Christiansen EH, Jensen LO, Thayssen P, et al.** Biolimus-eluting biodegradable polymer-coated stent versus durable polymer-coated sirolimus-eluting stent in unselected patients receiving percutaneous coronary intervention (SORT OUT V): a randomised non-inferiority trial. *Lancet.* 2013;381(9867):661–669.
25. **Paton NI, Kityo C, Hoppe A, et al.** Assessment of second-line antiretroviral regimens for HIV therapy in Africa. *N Engl J Med.* 2014;371(3):234–247.
26. **Kirchhof P, Andresen D, Bosch R, et al.** Short-term versus long-term antiarrhythmic drug treatment after cardioversion of atrial fibrillation (Flec-SL): a prospective, randomised, open-label, blinded endpoint assessment trial. *Lancet.* 2012;380(9838):238–246.
27. **Johnson P, Federico M, Kirkwood A, et al.** Adapted Treatment Guided by Interim PET-CT Scan in Advanced Hodgkin's Lymphoma. *N Engl J Med.* 2016;374(25):2419–2429.
28. **Radford J, Illidge T, Counsell N, et al.** Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. *N Engl J Med.* 2015;372(17):1598–1607.
29. **Park SJ, Ahn JM, Kim YH, et al.** Trial of everolimus-eluting stents or bypass surgery for coronary disease. *N Engl J Med.* 2015;372(13):1204–1212.
30. **Anderson CS, Robinson T, Lindley RI, et al.** Low-Dose versus Standard-Dose Intravenous Alteplase in Acute Ischemic Stroke. *N Engl J Med.* 2016;374(24):2313–2323.
31. **Boussier MG, Amarencu P, Chamorro A, et al.** Terutroban versus aspirin in patients with cerebral ischaemic events (PERFORM): a randomised, double-blind, parallel-group trial. *Lancet.* 2011;377(9782):2013–2022.
32. **Paul M, Bishara J, Yahav D, et al.** Trimethoprim-sulfamethoxazole versus vancomycin for severe infections caused by methicillin resistant *Staphylococcus aureus*: randomised controlled trial. *BMJ.* 2015;350:h2219.
33. **Hussain M, Tangen CM, Berry DL, et al.** Intermittent versus continuous androgen deprivation in prostate cancer. *N Engl J Med.* 2013;368(14):1314–1325.
34. **Behringer K, Goergen H, Hitz F, et al.** Omission of dacarbazine or bleomycin, or both, from the ABVD regimen in treatment of early-stage favourable Hodgkin's lymphoma (GHSG HD13): an open-label, randomised, non-inferiority trial. *Lancet.* 2015;385(9976):1418–1427.
35. **Kaul U, Bangalore S, Seth A, et al.** Paclitaxel-Eluting versus Everolimus-Eluting Coronary Stents in Diabetes. *N Engl J Med.* 2015;373(18):1709–1719.
36. **Bwakura-Dangarembizi M, Kendall L, Bakeera-Kitaka S, et al.** A randomized trial of prolonged co-trimoxazole in HIV-infected children in Africa. *N Engl J Med.* 2014;370(1):41–53.
37. **Jindani A, Harrison TS, Nunn AJ, et al.** High-dose rifampin with moxifloxacin for pulmonary tuberculosis. *N Engl J Med.* 2014;371(17):1599–1608.
38. **Mulvenna P, Nankivell M, Barton R, et al.** Dexamethasone and supportive care with or without whole brain radiotherapy in treating patients with non-small cell lung cancer with brain metastases unsuitable for resection or stereotactic radiotherapy (QUARTZ): results from a phase 3, non-inferiority, randomised trial. *Lancet.* 2016;388(10055):2004–2014.
39. **Aberegg S.** Reporting noninferiority trials. *JAMA.* 2013;309(15):1584–1585.
40. **Gladstone B, Vach W.** Analyzing noninferiority trials: it is time for advantage deficit assessment – an observational study of published noninferiority trials. *Open Access J Clin Trials.* 2015;7:11–21.
41. **Prieto L, Sacristán JA.** Problems and solutions in calculating quality-adjusted life years (QALYs). *Health Qual Life Outcomes.* 2003;1:80.