




OPEN ACCESS

# Non-inferiority trials in cardiology: what clinicians need to know

James T Leung,<sup>1</sup> Stephanie L Barnes,<sup>2</sup> Sidney T Lo,<sup>3</sup> Dominic Y Leung <sup>3</sup>

<sup>1</sup>Cardiology, Royal North Shore Hospital, Saint Leonards, New South Wales, Australia

<sup>2</sup>Neurology, Concord Repatriation General Hospital, Concord, New South Wales, Australia

<sup>3</sup>Cardiology, University of New South Wales, Liverpool Hospital, Liverpool, New South Wales, Australia

## Correspondence to

Professor Dominic Y Leung, Cardiology, University of New South Wales, Liverpool Hospital, Sydney, NSW 2052, Australia; d.leung@unsw.edu.au

Received 29 July 2019

Revised 13 September 2019

Accepted 19 September 2019

## ABSTRACT

Clinical trials traditionally aim to show a new treatment is superior to placebo or standard treatment, that is, superiority trials. There is an increasing number of trials demonstrating a new treatment is non-inferior to standard treatment. The hypotheses, design and interpretation of non-inferiority trials are different to superiority trials. Non-inferiority trials are designed with the notion that the new treatment offers advantages over standard treatment in certain important aspects. The non-inferior margin is a predetermined margin of difference between the new and standard treatment that is considered acceptable or tolerable for the new treatment to be considered 'similar' or 'not worse'. Both relative difference and absolute difference methods can be used to define the non-inferior margin. Sequential testing for non-inferiority and superiority is often performed. Non-inferiority trials may be necessary in situations where it is no longer ethical to test any new treatment against placebo. There are inherent assumptions in non-inferiority trials which may not be correct and which are not being tested. Successive non-inferiority trials may introduce less and less effective treatments even though these treatments may have been shown to be non-inferior. Furthermore, poor quality trials favour non-inferior results. Intention-to-treat analysis, the preferred way to analyse randomised trials, may favour non-inferiority. Both intention-to-treat and per-protocol analyses should be recommended in non-inferiority trials. Clinicians should be aware of the pitfalls of non-inferiority trials and not accept non-inferiority on face value. The focus should not be on the p values but on the effect size and confidence limits.

Clinical trials have flourished in medicine. Conducting clinical trials is one of the most important tools, allowing proper evaluation of treatment options. The results of clinical trials often form the backbone of practice guidelines which have permeated all subspecialties of medicine. With the potential to alter clinical practice, clinical trials have significant clinical, professional and public interests.

## SUPERIORITY TRIALS

Traditionally, investigators aim to demonstrate a new treatment is better than standard treatment or placebo in clinical trials. There have been landmark trials evaluating treatment of a wide range of conditions. The Stroke Prevention in Atrial Fibrillation (SPAF) trial<sup>1</sup> demonstrated that warfarin was more effective than placebo in stroke prevention in non-valvular atrial fibrillation (NVAf). The Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries (GUSTO) trial<sup>2</sup> showed that accelerated tissue

plasminogen activator (TPA) offered survival benefits over standard-dose TPA and streptokinase in ST-elevation myocardial infarction (STEMI). The Clopidogrel in Unstable Angina to Prevent Recurrent Events (CURE) trial<sup>3</sup> demonstrated the superiority of clopidogrel and aspirin over aspirin alone in non-STEMI. More recently, Further Cardiovascular Outcomes Research With PCSK9 Inhibition in Subjects With Elevated Risk (FOURIER) trial<sup>4</sup> showed that evolocumab was better than placebo in reducing major adverse cardiovascular events in patients with cardiovascular diseases.

The null hypothesis in superiority trials is that the new treatment is not better than standard treatment/placebo. The alternative hypothesis is that the new treatment is better than standard treatment/placebo. If results of the trial are statistically significant, the null hypothesis can be rejected and the alternative hypothesis accepted. If the results are not statistically significant, the null hypothesis cannot be rejected. The Clopidogrel for High Atherothrombotic Risk and Ischemic Stabilization, Management and Avoidance (CHARISMA) trial<sup>5</sup> hypothesised that aspirin plus clopidogrel was superior to aspirin in prevention of atherothrombotic events in stable cardiovascular disease. The trial showed a relative risk of 0.93 (95% CI 0.83 to 1.05, p=0.22). Therefore, the null hypothesis could not be rejected and the conclusion was aspirin and clopidogrel combination was not more effective than aspirin monotherapy.

## NON-INFERIORITY TRIALS

More recently, there is an increasing trend with clinical trials demonstrating a new treatment is non-inferior to standard treatment. Non-inferiority trials are designed with the notion that the new treatment offers advantages over standard treatment in certain important aspects. The Rivaroxaban Once Daily Oral Direct Factor Xa Inhibitor Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET AF) trial<sup>6</sup> showed rivaroxaban was non-inferior to warfarin in stroke prevention in NVAf. The Randomized Evaluation of Long Term Anticoagulant Therapy (RE-LY) trial<sup>7</sup> showed dabigatran was non-inferior to warfarin in prevention of stroke or systemic embolism in NVAf. In the Placement of Aortic Transcatheter Valves II (PARTNER II) trial,<sup>8</sup> transcatheter aortic valve implantation (TAVI) was similar to surgical aortic valve replacement in intermediate-risk patients. In patients with severe mitral regurgitation, MitraClip was shown to be non-inferior to surgery in freedom from death, surgery for valve dysfunction or severe mitral



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Leung JT, Barnes SL, Lo ST, et al. *Heart* Epub ahead of print: [please include Day Month Year]. doi:10.1136/heartjnl-2019-315772

**Table 1** Comparisons between superiority and non-inferiority trials

Characteristics	Superiority trials	Non-inferiority trials
<b>Null hypothesis</b>	New treatment not superior to standard treatment/ placebo	New treatment inferior to standard treatment
<b>Alternative hypothesis</b>	New treatment is better than standard treatment/ placebo	New treatment non-inferior to standard treatment
<b>Non-inferior margin</b>	Not applicable	Predetermined
<b>Sequential testing</b>	Not possible	Can be performed
<b>Significance level</b>	P<0.05 (two-sided)	P<0.025 (one-sided)
<b>Comparator</b>	Standard treatment or placebo	Standard treatment (seldom placebo)
<b>Possible outcomes</b>	1. New treatment superior (or inferior) to standard treatment/placebo 2. Inconclusive	See figure 1
<b>Subgroup analyses</b>	Possible	Possible

regurgitation in A Study of the Evalve Cardiovascular Valve Repair System - Endovascular Valve Edge-to-Edge REpair STudy II (EVEREST II) trial.<sup>9</sup> In Prospective randomized evaluation of the Watchman Left Atrial Appendage Closure device in patients with atrial fibrillation versus long-term warfarin therapy (PREVIAL) trial, Watchman device was non-inferior to warfarin in prevention of stroke or systemic embolism.<sup>10</sup> There has been an increasing number of non-inferiority trials in cardiology, especially in the area of new devices or technologies. A search term of 'non-inferiority' or 'non-inferior trials' in PubMed in 2018 returned 53 publications in 1995 and 789 in 2016.

Table 1 shows a comparison of important aspects between superiority and non-inferiority trials. Despite these differences,

it has been argued that the classification of trials as superiority or non-inferiority is arbitrary and somewhat artificial, especially with non-regulatory trials and when classifications of treatment groups as new or standard is ambiguous.<sup>11</sup>

### Hypothesis testing in non-inferiority trials

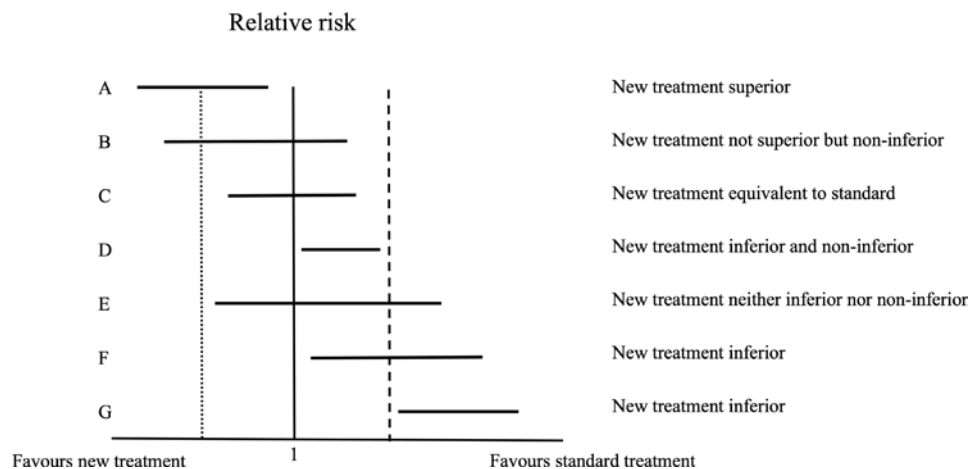
The null hypothesis in non-inferiority trials is that new treatment is inferior to standard treatment. The alternative hypothesis is the new treatment is non-inferior to standard treatment. If the results of the trial are statistically significant, the null hypothesis can be rejected and the alternative hypothesis that the new treatment is non-inferior to standard treatment can be accepted.

### Non-inferior margins

A non-inferior margin is specified in non-inferiority trials. The non-inferior margin is the predetermined margin of difference between the new and standard treatments. It represents how much worse the new treatment can be compared with standard treatment, yet still be considered 'similar' or 'not worse' than standard treatment. It is the maximal loss of efficacy of the standard treatment compared with the new treatment that can be tolerated in return for some perceived benefits of the new treatment. The non-inferior margin is defined relative to the established benefits of standard treatment over placebo. Figure 1 shows the hypothesis testing and possible outcomes of non-inferiority trials.

### Are non-inferiority trials necessary?

Non-inferiority trials may be necessary. For conditions like atrial fibrillation (AF), with proven effective therapies, it is no longer be ethical to test any new treatment against placebo. Therefore, many non-inferior trials used active controls as the comparator. In NVAf, any new oral anticoagulants would need to be tested



**Figure 1** Forest plot showing hypothesis testing and possible outcomes of non-inferiority trials. The solid vertical line indicates a relative risk of 1. The dashed vertical line indicates the predefined non-inferior margin. The dotted and dashed vertical lines indicate the equivalence margins (for equivalence trials). The horizontal lines represent the 95% CIs of the possible results of non-inferiority trials. A: the 95% CI does not overlap 1 and lies entirely on the left side of 1. Therefore, the new treatment is superior to standard treatment. B: the 95% CI overlaps 1 but the upper bound of the 95% CI does not exceed the predetermined non-inferior margin. Therefore, the new treatment is not superior but is non-inferior to standard treatment. C: the 95% CI overlaps 1 and lies entirely within the equivalence margins. Therefore, the new treatment is equivalent to standard treatment. D: the 95% CI does not overlap 1 and lies entirely on the right side of 1, Therefore, the new treatment is inferior to standard treatment. However, the upper bound of the 95% CI does not exceed the predetermined non-inferior margin. Therefore, the new treatment is also non-inferior to standard treatment. This scenario is unlikely but theoretically possible. E: the 95% CI overlaps 1 and the upper bound of the 95% CI exceeds the predetermined non-inferior margin. Therefore, the new treatment is neither inferior nor non-inferior to standard treatment. F and G: the 95% CIs lie entirely to the right side of 1 and the upper bound of the 95% CIs exceed the predetermined non-inferior margin. Therefore, the new treatment is inferior to standard treatment.

against warfarin as standard treatment for stroke prevention. Therefore, non-inferiority trial design with sequential testing would be suitable. Possible reasons for testing for non-inferiority for new treatment may be that the new treatment, while it may not be superior, may be cheaper, more convenient or may have fewer side effects than standard treatment.

The premise of non-inferiority trials is that the new treatment being tested is better than standard treatment in some important dimension. While the advantages of the new treatment over standard treatment are clear in some cases (eg, TAVI trials), it may not be as obvious in others. This may be illustrated by non-inferiority trials comparing different stent platforms like the Thin composite wire strut, durable polymer-coated (ResoluteOnyx) versus ultrathin cobalt–chromium strut, bioresorbable polymer-coated (Orsiro) drug-eluting stents in allcomers with coronary artery disease (BIONYX) trial.<sup>12</sup>

### Choosing the non-inferior margins

Two methods of choosing the non-inferior margin statistically can be used: relative risk difference and absolute risk difference.

#### Relative risk difference

A ratio of end point events on the new treatment to that on standard treatment is given as the non-inferior margin. The advantage of this method is that the event rate of the standard treatment does not need to be assumed. Non-inferiority is declared if the upper boundary of the 95% CI of the trial does not exceed that margin.

In trials of novel oral anticoagulants against warfarin in NVAF, the efficacy and confidence limits of warfarin over placebo were considered to determine the non-inferior margin. In general, the 50% rule is applied: at least 50% of the lower confidence limits (worst-case scenario) of the benefits of standard treatment over placebo is to be preserved. The choice of preserving 50% of the established treatment effects, although commonly practised, is empirical. In a meta-analysis of the six major trials of warfarin over placebo in non-valvular AF,<sup>13</sup> warfarin resulted in 62% (95% CI 48% to 72%) relative risk reduction of stroke or systemic embolism. In the worst-case scenario, warfarin resulted in 48% relative risk reduction (risk ratio 0.52). The risk ratio of adverse events on placebo over warfarin is the reciprocal of 0.52, that is, 1.92. To preserve 50% of 1.92 on the linear scale gives a non-inferior margin of 1.46 (the excess risk of placebo

over warfarin is 92%, half of that is 46%). On the log scale, preserving 50% of 1.92 is the square root of 1.92, that is, 1.38. In Apixaban for the Prevention of Stroke in Subjects With Atrial Fibrillation (ARISTOTLE),<sup>14</sup> the non-inferior margin was 1.44 (or 1.38 on the log scale). In Rivaroxaban Once Daily Oral Direct Factor Xa Inhibitor Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET AF), the non-inferior margin was 1.46.<sup>6</sup>

#### Absolute risk difference

Non-inferiority can be declared if the absolute difference in end points between the new and standard treatment is less than a predefined value. Unlike the relative risk method, this method entails an assumption on the event rate on standard treatment.

The actual event rate during the trial of the standard treatment is often lower than the assumed event rate (table 2). This will lead to a higher relative difference as the non-inferior margin and an underpowered trial, favouring non-inferiority. Table 2 shows the predefined non-inferior margins of several non-inferiority trials in cardiology with the assumed and observed event rates and the impact on the calculated relative risk difference.

#### Equivalence trials versus non-inferiority trials

In non-inferiority trials, investigators are interested in whether new treatment is non-inferior to standard treatment. Only the non-inferior margin to the right side of unity on the forest plot is specified. Therefore, the significance level is usually set as a one-sided p value of 0.025. In equivalence trials, investigators are interested in whether the new treatment is equivalent or 'the same as' standard treatment (figure 1). Therefore, margins on both sides of unity are specified. In equivalence trials, the significance level is set as a two-sided p value of 0.05. The COBALT trial hypothesised that double bolus TPA was equivalent to the accelerated regimen in STEMI.<sup>15</sup> The prespecified upper limit of absolute difference of 0.4% in 30-day mortality was exceeded. Therefore, the double bolus regimen was not shown to be equivalent to accelerated TPA with respect to 30-day mortality. Equivalence trials are seldom the focus now as non-inferiority trials with sequential testing are preferred.

In sequential testing, tests for non-inferiority are performed first. If non-inferiority is established, the results are then tested for superiority. Sequential testing is possible only if the trial is designed as a non-inferiority trial with a prespecified

**Table 2** The predefined non-inferior margins of non-inferiority trials with the assumed and observed event rates and the impact on the relative risk difference

Trial	Hypothesis	Non-inferior margin (absolute difference)	Assumed event rate on standard treatment	Calculated relative risk difference with assumed event rate	Observed event rate on standard treatment	Corresponding relative risk difference with observed event rate (with the predefined absolute difference non-inferior margin)
PARTNER IA <sup>27</sup>	TAVI non-inferior to surgical AVR in high-risk patients	7.5%	32%	1.23 (32%+7.5%)/32%	26.8%	1.28 (26.8%+7.5%)/26.8%
TREAT <sup>28</sup>	Ticagrelor non-inferior to clopidogrel after fibrinolysis for STEMI	1%	1.2%	1.83 (1.2%+1%)/1.2%	0.69%	2.44 (0.69%+1%)/0.69%
iFR SWEDEHEART <sup>29</sup>	iFR guided non-inferior to FFR-guided intervention	3.2%	8%	1.4 (8%+3.2%)/8%	6.1%	1.53 (6.1%+3.2%)/6.1%
PARTNER III <sup>30</sup>	TAVI non-inferior to surgical AVR in low-risk patients	6%	16.6%	1.36 (16.6%+6%)/16.6%	15.1%	1.4 (15.1%+6%)/15.1%
BIONYX <sup>12</sup>	Resolute Onyx stent non-inferior to Osiro stent	2.5%	6%	1.42 (6%+2.5%)/6%	4.7%	1.53 (4.7%+2.5%)/4.7%

The iFR SWEDEHEART trial showed an absolute difference between instantaneous wave-free ratio (iFR) and fractional flow reserve (FFR)-guided intervention of 0.7% (in favour of FFR-guided, 95% CI –1.5% to 2.8% with a relative risk of 1.12, 95% CI 0.79 to 1.58). This met the criteria for non-inferiority for iFR-guided intervention as the upper bound of the absolute difference did not exceed the predetermined non-inferior margin of 3.2%. However, if the relative risk method had been used, non-inferiority of iFR-guided intervention could not have been declared as the upper bound of the 95% CI of the relative risk was 1.58, exceeding 1.4. Even with the higher relative risk margin based on the observed event rate of the FFR arm (1.53), the actual results of the trial would not have met non-inferiority criteria.

AVR, aortic valve replacement; STEMI, ST-elevation myocardial infarction; TAVI, transcatheter aortic valve implantation.

non-inferior margin. Most non-inferiority trials now are tested for non-inferiority followed by testing for superiority. A to Z trial<sup>16</sup> evaluated enoxaparin against unfractionated heparin with a non-inferior margin of 1.144. It showed a non-significant hazard ratio (HR) of 0.88 favouring enoxaparin (95% CI 0.71 to 1.08). The upper boundary of the 95% CI did not exceed 1.144. Therefore, enoxaparin was declared non-inferior but not superior to unfractionated heparin. Sequential testing was similarly performed in other non-inferior trials including ROCKET AF and RE-LY. Caution should be exercised in interpreting p values of sequential testing, especially when p values of superiority and non-inferiority testing are reported side by side.

### Active versus placebo control

Most non-inferiority trials used active controls as the comparator. Placebo controls are seldom used in non-inferiority trials. However, placebo controls were used in some non-inferiority trials with different aims—examining safety of new treatments.

### Non-inferiority design examining safety

TECOS trial compared sitagliptin with placebo in type 2 diabetes.<sup>17</sup> The hypothesis was sitagliptin was non-inferior to placebo with respect to a primary combined cardiovascular outcome with a non-inferior margin of 1.3. The HR in TECOS trial was 0.98 in favour of sitagliptin (95% CI 0.89 to 1.08,  $p=0.68$ ). The upper boundary of the 95% CI did not exceed 1.3. Therefore, sitagliptin was declared non-inferior but not superior to placebo. The purpose of this non-inferiority trial with placebo control was to demonstrate the cardiovascular safety of new hypoglycaemic medications. The Food and Drug Administration (FDA) has mandated the demonstration of cardiovascular safety of new hypoglycaemic agents after the increased risks of death and myocardial infarction shown with rosiglitazone.<sup>18</sup> The non-inferior margin of 1.3 was stipulated by the FDA and was used in all such trials with the primary aim of demonstrating the cardiovascular safety of new hypoglycaemic medications. Liraglutide was compared with placebo in Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results (LEADER) trial<sup>19</sup> and empagliflozin was compared with placebo in the Empagliflozin Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients (EMPA-REG OUTCOME) trial,<sup>20</sup> both with non-inferior margin of 1.3. Unlike the TECOS trial, liraglutide and empagliflozin were shown to be non-inferior and superior to placebo with respect to cardiovascular outcomes by sequential testing.

### Assumptions of non-inferiority trials

Two inherent assumptions in non-inferiority trials are not necessarily tested in the trial and may not be correct. The first is that new treatment offers advantages over standard treatment. The new treatment may be cheaper, more convenient or easier to administer, more readily available, less invasive or may have fewer side effects. These properties are usually not tested in the trial.

The second assumption is that the new treatment is superior to placebo. This assumption is not tested unless a placebo arm is included. While a placebo control may be used in non-inferiority trials, the purpose of having placebo control is usually not to demonstrate the superior therapeutic effects of the new treatment but to demonstrate its safety. While the new treatment may be demonstrated to be non-inferior to the standard treatment in the trial, this does not necessarily mean it is superior to placebo. This is particularly the case if the standard treatment has only a

modest therapeutic advantage over placebo and/or a wide, unjustified non-inferior margin is chosen.

### Risks of successive non-inferiority trials

The risks of successive non-inferiority trials are unique to their design. If a new treatment is shown to be non-inferior to standard treatment, all it means is that the new treatment is not worse than the standard treatment by the predetermined non-inferior margin. The new treatment may subsequently replace the old standard treatment and be accepted as the new standard treatment.

If a newer treatment becomes available in the future, that newer treatment is likely to be compared with the new 'standard' treatment and not the old standard treatment in further non-inferiority trials. While the newer treatment may be non-inferior to the new standard treatment, it may no longer be non-inferior to the original standard treatment. In this way, less and less effective treatments may be introduced in successive non-inferiority trials.

### Intention-to-treat analyses

Intention-to-treat analyses may favour non-inferiority results. Protocol violations may dilute any potential differences between the treatment arms, favouring non-inferiority results. Therefore, it is preferable for non-inferior trials to be analysed according to both the intention-to-treat and per-protocol approach. Non-inferiority may be declared only if both analyses support the same conclusion. ROCKET AF and RE-LY trials provided both intention-to-treat and per-protocol analyses while PARTNER III and Evolut low-risk trial<sup>21</sup> reported per-protocol analysis.

### Impact of the quality of the non-inferiority trials on results

The quality of the design and execution of clinical trials have a significant impact on the results. Having adequate statistical power, well-defined and high-quality end points, adequate follow-up duration, low loss-to-follow-up and low protocol-violation rates are critical for the success of superiority trials.

Low-quality trials, however, may favour non-inferior results. Poorly defined or soft end points, inadequate statistical power and follow-up duration, high protocol violation and loss-to-follow-up may dilute any potential differences between the new and standard treatment and therefore favour non-inferiority. Carefully choosing the study population is of critical importance in any clinical trial. It is perhaps more important in non-inferiority trials. Choosing low-risk population in testing for superiority of a treatment over placebo will likely lead to non-significant results. In contrast, having low-risk population in non-inferiority trials may favour non-inferiority. Furthermore, having more stringent criteria for statistical significance, that is, one-sided p value of 0.025 is considered desirable for non-inferior trials compared with the usual  $\alpha$  of 0.05. In this respect, practices varied. Trials like ROCKET AF, ARISTOLE, RE-LY, LEADER and EMPA-REG OUTCOME trials used  $\alpha$  of one-sided p value of 0.025, whereas EVEREST II and PARTNER IA trial used  $\alpha$  of one-sided p value of 0.05.

The inclusion of soft end points leads to a higher event rate and therefore allows a more generous non-inferior margin to be chosen. A higher non-inferior margin can be chosen by the relative risk difference method as it is easier to preserve half of the benefits of the standard treatment over placebo due to the higher event rate. Furthermore, assuming a higher event rate may favour non-inferiority.

### Reporting of non-inferiority trials

Inadequacies were common in the reporting of non-inferiority trials in the literature. Common deficiencies include



inappropriate use of non-inferiority design, inappropriate claims of non-inferiority or equivalence with non-significant results in superiority trials, failure to provide rationale for the non-inferiority design or non-inferior margins.

In a review of 88 equivalence trials from 1992 to 1996, equivalence was inappropriately declared from non-significant results of superiority testing in two-thirds.<sup>22</sup> The GUSTO III trial compared reteplase with accelerated TPA in STEMI.<sup>23</sup> The trial was designed as a superiority trial to detect a 20% relative risk reduction in mortality with reteplase. The trial showed a non-significant increase in mortality of 0.23% (95% CI -1.11% to 0.66%) with reteplase. The trial concluded that reteplase and accelerated TPA were similar in their effects on 30-day mortality. Equivalence in this case was inappropriately claimed from a non-significant result in what was designed as a superiority trial.

Many non-inferiority trials failed to report a predetermined non-inferior margin.<sup>22</sup> Even when a non-inferior margin was specified, less than half reported how the margin was determined.<sup>24,25</sup> Of the 232 non-inferiority trials, only 17 (7.3%) took into account preserving effects of standard treatment, with 15 applying the 50% rule.<sup>24</sup> The trial results were analysed according to the intention-to-treat principle in 34.9% and per-protocol analysis in 19.4%. Only 41.8% of the trials analysed results by both methods.<sup>25</sup>

The Consolidated Standards of Reporting Trials (CONSORT) group provided an extension of the 2010 CONSORT statements in 2012 for reporting non-inferiority and equivalence trials.<sup>26</sup> The extension of the checklist includes identification of the trial as a non-inferiority randomised trial in the title, providing the rationale for the non-inferior design, specifying the non-inferior margin and its rationale and specifying the non-inferiority outcomes.

### Non-inferiority trials: what the clinician needs to know

The interpretation of non-inferiority trials is more complex than that of superiority trials. Correct interpretation involves carefully examining the design, the execution as well as the analysis and results of the trials. Is the non-inferiority trial design appropriate? What is the primary non-inferior hypothesis? What is the non-inferior margin and how is it chosen? These must be stated in all non-inferiority trials. The non-inferior margin must be specified with the rationale clearly specified. Is the study population representative of the condition being studied? An important aspect is to look at the event rate of the standard treatment, which should be comparable to historical results in previous superiority trials of the standard treatment over placebo. This will allow the reader to judge if suitably at-risk patients are recruited in the non-inferiority trial. It is also important to examine the quality of the execution of the trial. Poor quality trials favour non-inferior results. Non-inferiority trials should be analysed according to both intention-to-treat and per-protocol approaches and the results of both analyses should be in concordance. The point estimate and the CIs of the non-inferior trial should be noted as well, not just the p values showing non-inferiority. This will allow the clinician to assess how close the results of the study are to the non-inferior margin.

The use of non-inferiority trials is likely to increase. Clinicians should not accept the term 'non-inferior' at face value. It is important to be aware that 'non-inferior' does not mean 'equivalent' or 'similar' or 'as good as'. Non-inferiority means the new treatment is not unacceptably worse than standard treatment.

**Correction notice** Since this article was first published online, figure 1 has been replaced. Further details on the changes can be found in the eletters section of the paper. These eletters can be found in the responses tab of the online article.

**Contributors** All authors have contributed to the writing, editing, revision and final approval of the manuscript.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

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

### ORCID iD

Dominic Y Leung <http://orcid.org/0000-0002-5626-7236>

### REFERENCES

- 1 Stroke prevention in atrial fibrillation study. final results. *Circulation* 1991;84:527–39.
- 2 An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. *N Engl J Med Overseas Ed* 1993;329:673–82.
- 3 Yusuf S, Zhao F, Mehta SR, *et al*. Effects of clopidogrel in addition to aspirin in patients with acute coronary syndromes without ST-segment elevation. *N Engl J Med* 2001;345:494–502.
- 4 Sabatine MS, Giugliano RP, Keech AC, *et al*. Evolocumab and clinical outcomes in patients with cardiovascular disease. *N Engl J Med* 2017;376:1713–22.
- 5 Bhatt DL, KAA F, Hacke W, *et al*. Clopidogrel and aspirin versus aspirin alone for the prevention of atherothrombotic events. *New Eng J Med* 2006;354:1706–17.
- 6 Patel MR, Mahaffey KW, Garg J, *et al*. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med* 2011;365:883–91.
- 7 Connolly SJ, Ezekowitz MD, Yusuf S, *et al*. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009;361:1139–51.
- 8 Leon MB, Smith CR, Mack MJ, *et al*. Transcatheter or surgical aortic-valve replacement in intermediate-risk patients. *N Engl J Med* 2016;374:1609–20.
- 9 Feldman T, Foster E, Glower DD, *et al*. Percutaneous repair or surgery for mitral regurgitation. *N Engl J Med* 2011;364:1395–406.
- 10 Holmes DR, Kar S, Price MJ, *et al*. Prospective randomized evaluation of the Watchman left atrial appendage closure device in patients with atrial fibrillation versus long-term warfarin therapy. *J Am Coll Cardiol* 2014;64:1–12.
- 11 Dunn DT, Copas AJ, Brocklehurst P. Superiority and non-inferiority: two sides of the same coin? *Trials* 2018;19:499.
- 12 von Birgelen C, Zocca P, Buiten RA, *et al*. Thin composite wire strut, durable polymer-coated (Resolute Onyx) versus ultrathin cobalt–chromium strut, bioresorbable polymer-coated (Orsiro) drug-eluting stents in allcomers with coronary artery disease (BIONYX): an international, single-blind, randomised non-inferiority trial. *The Lancet* 2018;392:1235–45.
- 13 Hart RG, Benavente O, McBride R, *et al*. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation. *Ann Intern Med* 1999;131:492–501.
- 14 Granger CB, Alexander JH, McMurray JJV, *et al*. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011;365:981–92.
- 15 Continuous Infusion versus Double-Bolus Administration of Alteplase (COBALT) Investigators. Continuous infusion versus double-bolus administration of alteplase I. A comparison of continuous infusion of alteplase with double-bolus administration for acute myocardial infarction. *N Engl J Med* 1997;337:1124–30.
- 16 Blazing MA *et al*. Safety and Efficacy of Enoxaparin vs Unfractionated Heparin in Patients With Non-ST-Segment Elevation Acute Coronary Syndromes Who Receive Tirofiban and Aspirin<SUBTITLE>A Randomized Controlled Trial</SUBTITLE>&g. *JAMA* 2004;292:55–64.
- 17 Green JB, Bethel MA, Armstrong PW, *et al*. Effect of sitagliptin on cardiovascular outcomes in type 2 diabetes. *N Engl J Med* 2015;373:232–42.
- 18 Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med* 2007;356:2457–71.
- 19 Marso SP, Daniels GH, Brown-Frandsen K, *et al*. Liraglutide and cardiovascular outcomes in type 2 diabetes. *N Engl J Med* 2016;375:311–22.
- 20 Zinman B, Wanner C, Lachin JM, *et al*. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med* 2015;373:2117–28.
- 21 Popma JJ, Deeb GM, Yakubov SJ, *et al*. Transcatheter aortic-valve replacement with a self-expanding valve in low-risk patients. *N Engl J Med* 2019;380:1706–15.
- 22 Greene WL, Concato J, Feinstein AR. Claims of equivalence in medical research: are they supported by the evidence? *Ann Intern Med* 2000;132:715–22.

- 23 Global use of strategies to open occluded coronary arteries I. A comparison of reteplase with alteplase for acute myocardial infarction. *N Engl J Med* 1997;337:1118–23.
- 24 Wangge G, Klungel OH, Roes KCB, *et al.* Interpretation and inference in noninferiority randomized controlled trials in drug research. *Clin Pharmacol Ther* 2010;88:420–3.
- 25 Wangge G, Klungel OH, Roes KCB, *et al.* Room for improvement in conducting and reporting non-inferiority randomized controlled trials on drugs: a systematic review. *PLoS One* 2010;5:e13550.
- 26 Piaggio G, Elbourne DR, Pocock SJ, *et al.* Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012;308:2594–604.
- 27 Smith CR, Leon MB, Mack MJ, *et al.* Transcatheter versus surgical aortic-valve replacement in high-risk patients. *N Engl J Med* 2011;364:2187–98.
- 28 Berwanger O, Nicolau JC, Carvalho AC, *et al.* Ticagrelor vs clopidogrel after fibrinolytic therapy in patients with ST-elevation myocardial infarction: a randomized clinical trial. *JAMA Cardiol* 2018;3:391–9.
- 29 Götberg M, Christiansen EH, Gudmundsdottir IJ, *et al.* Instantaneous Wave-free ratio versus fractional flow reserve to guide PCI. *N Engl J Med* 2017;376:1813–23.
- 30 Mack MJ, Leon MB, Thourani VH, *et al.* Transcatheter aortic-valve replacement with a Balloon-Expandable valve in low-risk patients. *N Engl J Med* 2019;380:1695–705.