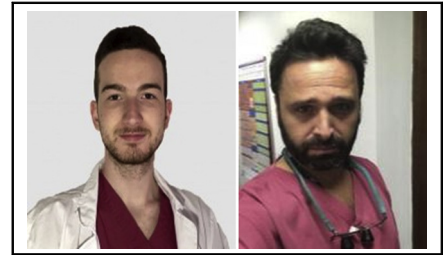


See Article page 2119.



Commentary: Noninferiority trial: The devil is...

Arnaldo Dimagli, MD, and Umberto Benedetto, PhD



Arnaldo Dimagli, MD, and Umberto Benedetto, PhD

There are 3 main types of clinical trials based on the statistical hypothesis tested. Superiority trials compare 2 treatments to determine which is better, equivalence trials test whether a new treatment is as effective as the standard treatment, and noninferiority trials aim to demonstrate that a new experimental treatment is not unacceptably worse than the standard treatment. These latter studies are being more and more embraced¹ because they allow investigators to study new treatments that can bring an advantage in terms of costs, patient compliance to medication, fewer adverse effects, or greater availability.²

Falk and Friede³ focus on the main aspects of noninferiority trials with prodigious attention to details. The authors describe the critical step of choosing a noninferiority margin, that is the value that the effect measure of the experimental treatment must not exceed when compared with the standard treatment to prove its noninferiority. The noninferiority margin should be oriented from both a statistical and a clinical point of view and should be prospectively calculated and clearly reported in the protocol of the study.⁴ Once the noninferiority margin is set, the trial null hypothesis will test that the treatment difference between the standard and the experimental strategy is equal to or greater than the assumed margin whereas the alternative hypothesis will state that the efficacy value of the standard treatment and the experimental treatment differ by no more than the margin. It is important to acknowledge the different statistical hypotheses that are being tested in superiority and inferiority trials to interpret results accordingly. In a superiority trial, the acceptance of the null hypothesis—the new treatment is not superior to the standard one—does not allow us to draw a conclusion

CENTRAL MESSAGE

Investigators of noninferiority trials need thorough understanding of statistical and technical peculiarities for their results to have valuable clinical influence.

of noninferiority because this would lead to a biased estimate. Conversely, in a noninferiority trial, when the null hypothesis is rejected, we assume that the new treatment is noninferior to the comparator and that the 2 treatments only differ by clinically irrelevant effects.

But what does clinically irrelevant mean for patients? A practical example may help. Would you accept to buy a house located in a less-safe neighborhood just because it is less expensive than another one in a safer area? If the difference in the crime rates was only 0.1%, it would be reasonable to save 10% on the cost of the house, but if the savings were 0.1% and the crimes rate were 10% higher it would be hard to fairly choose the less-expensive house. This applies to patients in our daily practice. As suggested by the European Medicines Agency,⁴ there are areas where mortality or other irreversible morbidities (eg, stroke) may not allow any margin of inferiority. It would be extremely difficult to ethically accept an excess rate in deaths from a new treatment just because it is less expensive or with minor adverse effects. For example, a QID drug A has proved over decades to be safe and efficient in preventing death, stroke, and myocardial infarction. A new QD treatment B would be positively accepted for its anticipated better compliance to treatment as long as it will not cause more deaths, strokes, and myocardial infarctions than drug A. Here comes an intriguing paradox: if we assume that drug B will increase the compliance to the treatment, we can reasonably anticipate that there will be a relevant advantage that translates into a superior

From the Bristol Heart Institute, University of Bristol, Bristol, United Kingdom.

Disclosures: The authors reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

Received for publication April 9, 2020; accepted for publication April 10, 2020; available ahead of print April 23, 2020.

Address for reprints: Umberto Benedetto, PhD, Bristol Heart Institute, University of Bristol, Bristol Royal Infirmary, Upper Maudlin St, BS2 8HW, Bristol, United Kingdom (E-mail: umberto.benedetto@bristol.ac.uk).

J Thorac Cardiovasc Surg 2021;161:2124-5

0022-5223/\$36.00

Copyright © 2020 by The American Association for Thoracic Surgery

<https://doi.org/10.1016/j.jtcvs.2020.04.043>

outcome and goes beyond the pure evidence of having only a more comfortable regimen for a patient. So noninferiority is being tested when superiority could be proved.

Noninferiority trials can represent a rigorous method to widen treatment options in proper settings, but physicians should carefully read and understand the deepest statistical and technical pieces of them to verify that a proper noninferiority margin was chosen and assess the extent of clinically irrelevant events.

See Article page 2119.



Commentary: How to overcome an inferiority complex

Paul Kurlansky, MD

Noninferiority trials require careful execution and interpretation. For a variety of reasons, this format has increased markedly in recent years, with a PubMed search revealing 1 study each in 1997 and 1998, with an acceleration to greater than 500 in 2015¹ (Figure 1). Using “noninferiority” as a search term in [ClinicalTrials.gov](https://clinicaltrials.gov) currently yields 2441 studies.² Therefore, the expert review by Falk and Friede³ in this issue of the *Journal* is both timely and welcome. At the risk of redundancy, it is worthwhile to review several of the pertinent points that they have raised (although their guide to many of the technical aspects of designing and interpreting a noninferiority trial will prove invaluable researchers and clinicians), as well as to explore how the current literature meets many of the standards that they propose.

Interestingly enough, if I dust off my trusty (if ancient) *Merriam-Webster's Collegiate Dictionary* (or even my son's more recent edition), I discover that “noninferior” and “noninferiority” are not recorded as words in the

From the Division of Cardiac Surgery, Department of Surgery, Columbia University, New York, NY.

Disclosures: The author reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

Received for publication April 28, 2020; revisions received April 28, 2020; accepted for publication April 29, 2020; available ahead of print May 16, 2020.

Address for reprints: Paul Kurlansky, MD, Black Building 210, 650 West 168th St, New York, NY 10032 (E-mail: Pk2245@cumc.columbia.edu).

J Thorac Cardiovasc Surg 2021;161:2125-9

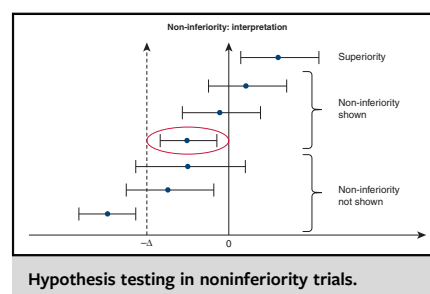
0022-5223/\$36.00

Copyright © 2020 by The American Association for Thoracic Surgery

<https://doi.org/10.1016/j.jtcvs.2020.04.153>

References

1. 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.
2. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA*. 2012;308:2594-604.
3. Falk V, Friede T. Noninferiority trials: what's clinically (ir) relevant? *J Thorac Cardiovasc Surg*. 2021;161:2119-23.
4. European Medicines Agency. Guideline on the choice of the non-inferiority margin. Available at: www.ema.europa.eu/pdfs/human/ewp/215899en.pdf. Accessed April 6, 2020.



CENTRAL MESSAGE

Noninferiority trials have increased in prevalence and importance, yet require specialized understanding and care to properly conduct and interpret.

English language. Perhaps it is this very “nonintuitive” nature that makes these studies so challenging to perform and interpret appropriately. Unlike the classic superiority trial, where treatment A is compared with placebo or treatment B to determine if one is better (however defined) than the other, a noninferiority trial seeks to determine if A is no worse than B. Therefore, the null hypothesis, rather than being that there is no difference between A and B, is actually that A is worse than B. This shift turns many things somewhat upside down—demonstration of a difference may actually lead to acceptance rather than rejection of the null hypothesis. Why do this? Aside from the more cynical perspective that this approach paves the way for “me too” drugs that pharmaceutical companies develop for market share but that add little or no value to those currently