

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Udayasiri, DK;Skandarajah, A;Hayes, IP

Title:

The inferiority of noninferiority trials

Date:

2020-03-01

Citation:

Udayasiri, D. K., Skandarajah, A. & Hayes, I. P. (2020). The inferiority of noninferiority trials. ANZ Journal of Surgery, 90 (3), pp.203-205. <https://doi.org/10.1111/ans.15592>.

Persistent Link:

<https://hdl.handle.net/11343/275482>

Title Page

Title

The Inferiority of Noninferiority Trials.

Short running head

Noninferiority Trials

Authors

Dilshan K. Udayasiri

Degrees: M.B.B.S., F.R.A.C.S.

Academic title: Dr

Main appointment:

Colorectal research fellow

Institution/department:

Colorectal Surgical Unit, The Royal Melbourne Hospital, Victoria, Australia

Department of General Surgical Specialties, The Royal Melbourne Hospital, Victoria,
Australia

Department of Surgery, The University of Melbourne, Victoria, Australia

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1111/ans.15592](#)

Anita Skandarajah

Degrees: M.B.B.S., M.D., F.R.A.C.S., F.A.C.S

Academic title: Dr

Main Appointment:

Breast and Endocrine Surgeon

Co Supervisor for research

Institution/department:

Department of General Surgical Specialties, The Royal Melbourne Hospital, Victoria,
Australia

Department of Surgery, The University of Melbourne, Victoria, Australia

Ian P. Hayes

Degrees: M.B.B.S., M.S., M.Epi., F.R.A.C.S., F.R.C.S. (Gen-Surg)

Academic title: Clinical Associate Professor

Main Appointment:

Chief Supervisor of Research

Head of Colorectal Surgical Unit

Colorectal Surgeon

Institution/department:

Colorectal Surgical Unit, The Royal Melbourne Hospital, Victoria, Australia

Department of General Surgical Specialties, The Royal Melbourne Hospital, Victoria,
Australia

Department of Surgery, The University of Melbourne, Victoria, Australia

Author responsible for correspondence about the manuscript

Dr Dilshan Udayasiri

9 Cook Street, Ivanhoe, Victoria, Australia, postcode 3079

Telephone: +61431059868

Email: dilshan.udayasiri@gmail.com

Sources of support

Condon Colonic Disorder Scholarship, Medicine, Dentistry and Health Sciences, The University of Melbourne.

Health Technology Assessment Scholarship, The Royal Australian College of Surgeons

Podium or Poster Presentations

Nil

Word count of text: 999

Number of tables: 4

Disclaimers

There are no conflicts of interest in this article.

Noninferiority trials are increasingly popular in surgical literature. They aim to demonstrate that a new treatment is no worse than an existing treatment but may have added benefits. Their design is, in many ways, counterintuitive and prone to bias. In this article we will highlight the design features of noninferiority trials by contrasting them to other studies. We will then explore the biases inherent in their design.

A key to understanding noninferiority trials is that their null hypothesis is structured differently to other study designs. In most comparative studies, the null hypothesis is that there is no difference between the intervention and control. The purpose of statistical analysis is to search for evidence against the null hypothesis. The strength of this evidence, represented by the p-value, dictates whether we can accept or reject the null hypothesis. A low p-value provides evidence against the null hypothesis whereas a high p-value fails to do so.

In most two-sided comparative studies, the result can show either a difference between intervention and control or no difference. If a difference is found, its size and direction indicates whether the intervention is better or worse than control.

One-sided studies, in contrast, are more restrictive. We will commence the discussion with superiority trials, which can be two sided, but for the purposes of this explanation we will restrict our definition of superiority trials to those which are one-sided. In the case of a one-sided superiority trial, the aim is to investigate whether an intervention is superior to control by a specified margin (superiority margin). The null hypothesis is that the intervention is NOT SUPERIOR to control. Not-superior in this case includes equivalence and inferiority

(Figure 1) A two-sided study, on the other hand, would be able to differentiate between equivalence and inferiority. A one-sided superiority trial can only analyse if an intervention is superior or not-superior to control. To continue this line of reasoning we feel it is helpful to conceptualise the opposite of a one-sided superiority study, which would be an inferiority study (never performed). This hypothetical one-sided study would investigate whether an intervention is inferior to control. The null hypothesis is that the intervention is NOT INFERIOR to control. Not-inferior includes equivalence and superiority (Figure 2).

In a noninferiority trial however, there is an important change to the structure of the null hypothesis. The null hypothesis here is that the intervention is INFERIOR to control. Thus, the null hypothesis of a noninferiority trial does not include equivalence, in contrast to the null hypotheses of superiority and inferiority trials where equivalence is included (Figure 3). If there is evidence against the null hypothesis of a noninferiority trial, then the intervention is NONINFERIOR to control which could either mean the intervention is equivalent or superior to control.

A common misconception with a noninferiority trial is that, if the result is not significant ($P > 0.05$), the intervention is equivalent to control; this is not true. In fact, if there is no evidence against the null hypothesis, the intervention in a noninferiority trial must be inferior to control (Figure 4).

It is important to understand the biases inherent in a noninferiority study design.

Failure to reject a null hypothesis when there is a real difference between intervention and control is a Type 2 error. When the difference between treatment groups is small, a large

number of patients is required in each group to demonstrate this difference. This is true for any study where the null hypothesis includes equivalence (two-sided study, superiority study, inferiority study). In a small noninferiority study where the intervention is truly inferior to control, it may be difficult to demonstrate this difference and equivalence/superiority is more likely to be found; a false rejection of the null hypothesis. Thus, in small noninferiority studies, an intervention may be found to be equivalent to control when it is, in fact, inferior.

The noninferiority trial design aims to show that a new intervention is no worse than control by a prespecified margin, defined by the symbol Δ (delta). This margin is used in a manner analogous to the 95% confidence interval (CI) used in other trials. The size of the selected margin is critical: too narrow and a truly noninferior intervention may be found to be inferior, too wide and a truly inferior intervention may be found to be noninferior.

The ideal way of selecting Δ is using *Historical Evidence* where an active treatment was compared to placebo (traditionally used in drug trials). This active treatment then becomes the control arm in a new noninferiority trial with a novel active treatment. The extent to which the original active treatment was superior to placebo becomes the noninferiority margin of the new trial. In this way, the trial aims to demonstrate that the new intervention is no worse than control by the same amount as control is superior to placebo. To ensure that Δ is as conservative as possible, the lower boundary of the 95%CI of the control vs placebo trial is used as Δ for the noninferiority trial comparing the novel intervention against control.

In surgical noninferiority trials, the Historical Evidence model is rarely used as there are very few studies comparing a surgical procedure to placebo. Therefore, other, less objective

methods of selecting Δ are required. Expert opinion or margins previously used in other similar studies determine the margin below which the new intervention is no worse than control. In contrast to the Historical Evidence technique, this method of margin selection is prone to bias.

Noninferiority studies are becoming increasingly popular in surgical literature and therefore it is important to understand their design, especially the null hypothesis being that the new intervention is inferior to the control. Their design favours rejecting a true null hypothesis in studies with smaller numbers; the opposite to other study designs. How a noninferiority margin is chosen should be stated. In surgical literature, margin choice is commonly based on professional opinion which is prone to bias and larger margins may falsely reject a true null hypothesis.

Figure Legends

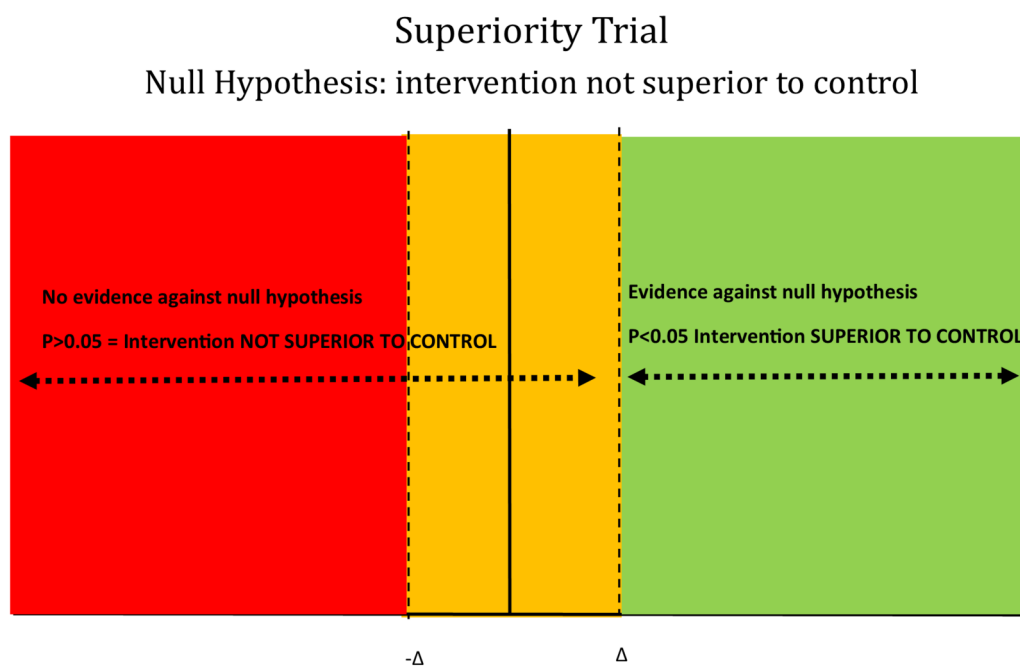


Figure 1

- Zone of inferiority to Control
 - Zone of Equivalence to Control, between $-\Delta$ and Δ .
 - Zone of Superiority to Control
- Δ = superiority margin

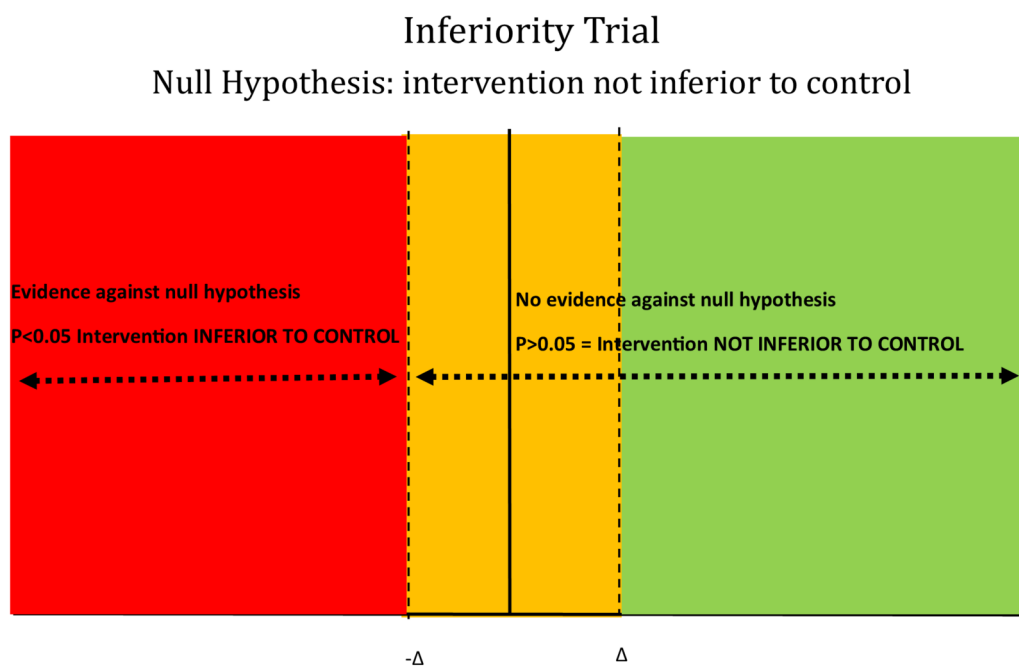


Figure 2

- Zone of inferiority to Control
- Zone of Equivalence to Control, between $-\Delta$ and Δ .
- Zone of Superiority to Control
- $-\Delta$ = inferiority margin

Noninferiority Trial

Null Hypothesis: intervention inferior to control

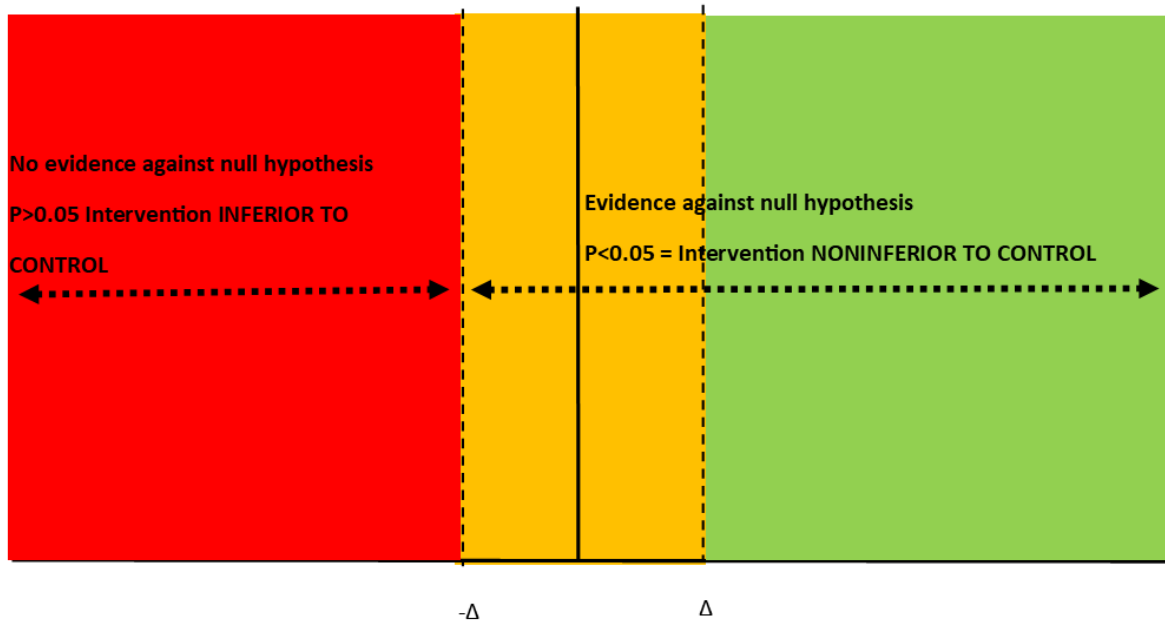


Figure 3

- Zone of inferiority to Control
- Zone of Equivalence to Control, between $-\Delta$ and Δ .
- Zone of Superiority to Control
- $-\Delta$ = noninferiority margin

Noninferiority study

Null hypothesis: intervention inferior to control

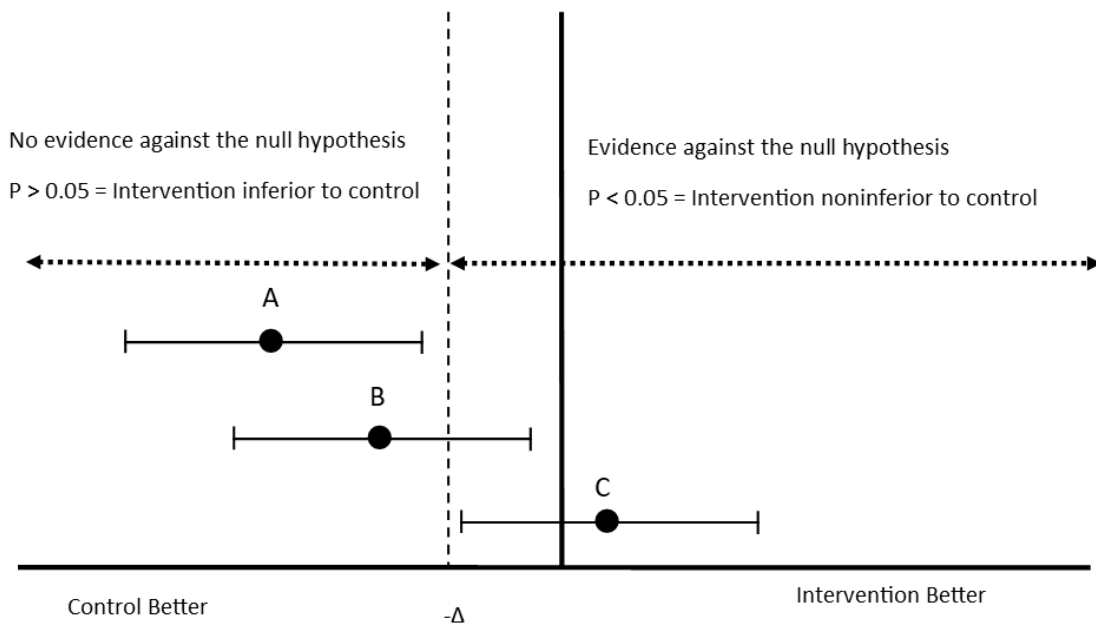


Figure 4

- A. Intervention inferior to control.
- B. Intervention inferior to control as 95%CI crosses delta and $P > 0.05$ therefore no convincing evidence against the null hypothesis.
- C. Intervention noninferior to control.
- $-\Delta$: noninferiority margin.