Noninferiority trials: clinical understandings and misunderstandings

John H Powers1 & Thomas R Fleming*2

In many clinical settings for the treatment or prevention of disease, currently available interventions provide clinically meaningful benefits by decreasing irreversible morbidity or mortality. An important example is the setting of treatment for pneumonia, where some antibiotics provide substantive benefits by meaningfully reducing mortality in addition to improving symptoms of cough, breathlessness and chest pain. Even though existing standard interventions (hereafter called ‘Standard’) provide important clinical benefits in clinical settings such as pneumonia, patients and caregivers may be interested in new interventions that would essentially retain the efficacy of Standard while being substantially better in terms of safety, convenience of administration, or cost.

When the goal is to replace an existing Standard that provides clinically meaningful effects on measures of irreversible morbidity or mortality, there is ethical motivation to use Standard as the control in a randomized trial evaluating the new intervention. While it would be preferable in such a trial to establish the new intervention to have superior efficacy, it may be sufficient to rule out that its efficacy is unacceptably worse than that of Standard. These are called noninferiority (NI) trials and have been discussed extensively in the literature [1–12,101]. By design, the NI trial requires specification of the minimum threshold constituting an unacceptable loss of efficacy. This threshold is called the NI margin.

The formulation of the NI margin is often controversial. Sponsors interested in conducting smaller trials and increasing the likelihood of achieving ‘positive’ results prefer large margins. However, to avoid exposing patients to meaningfully less-effective new interventions, there should be rigorous scientific justification for the NI margin.

An inherent weakness of the NI trial design is that there is not a placebo ‘anchor’. If results of the NI trial reveal the new intervention and existing standard intervention have similar efficacy, are these regimens similarly effective or ineffective in that setting?

Keywords: active control • informed consent • noninferiority margin • preservation of effect
A new experimental antibiotic that truly is ineffective in all patients may mistakenly be judged to be effective if it is evaluated in a noninferiority comparison with Standard that is conducted in only young patients at low risk for major morbidity or mortality, if the Standard is ineffective or has much less effect on the absolute risk of death in such low risk patients [12].

Another key consideration in the choice of the NI margin is ensuring a substantial fraction of the effect of Standard is preserved by an alternative regimen, especially in settings where it would be unethical to deprive control patients access to Standard due to its meaningful effect on risks of irreversible morbidity or mortality [12,13,102].

**Clinical understandings & misunderstandings: some important issues**

There is confusion about the purpose of NI trials and, in turn, about their design, application and interpretation. The term 'NI' is itself confusing as this implies that the conclusion of a positive NI trial is that the new intervention is 'not worse' than Standard. However, an intervention may be statistically inferior to Standard and still meet a definition of NI specified for that trial. A premise of ‘NI’ is that differences smaller than the NI margin are not clinically consequential. Hence, by ruling out the prespecified NI margin, the estimated difference between the new intervention and Standard is then statistically inconsistent with any true levels of efficacy loss that would be clinically consequential.

Unfortunately, investigators, clinicians and patients often believe that an estimate of ‘no difference’ in NI trials translates into equality between the new intervention and Standard such that the regimens are entirely interchangeable. This misperception leads to several other consequences. The medical literature shows that many trials declared to show NI are failed superiority trials with no prespecified NI margin [14]. The reporting of NI trials is generally poor [15].

An intervention that ‘works’ may not be similarly effective or have a similar risk–benefit assessment under all conditions and in all types of patients. This fact provides challenges both in the justification of the NI margin as well as in the interpretation of results of a NI trial. Regarding the justification of the NI margin, international guidance indicates the historical evidence used to estimate the effect of Standard in the NI trial needs to come from settings that match the NI trial’s definition of disease setting, patient population, prior and concomitant medication, outcome and timing of analysis [2]. Regarding interpretation of results of a NI trial, a positive NI trial in one population does not allow one to conclude superiority (against placebo) in another clinical setting. However, in some fields such as infectious diseases, drug sponsors commonly attempt to claim that a new intervention is superior in an unstudied population of patients based on the results of a NI trial conducted in another setting. For example, establishing NI in patients with susceptible disease does not establish superiority in an unstudied population with resistant disease, given that patients with resistant pathogens are often older, sicker and have more co-morbidities [16].

Since NI trials are actively controlled, some consider them inherently 'more ethical' because they do not expose patients to a placebo. However, NI trials raise ethical questions of their own. It is unethical to conduct a poorly designed NI trial that exposes patients to potential harm without benefit for themselves or for society. One criterion for a properly designed and ethical NI trial is a reliable understanding of how much loss of effect is still 'clinically acceptable' to a patient. This understanding should be obtained by querying patients rather than by interviewing only clinicians. Furthermore, when Standard previously has been shown to meaningfully reduce risks of mortality or major morbidity, the goal of a NI trial should be to determine whether the new intervention preserves an adequate amount of Standard's effect, not solely to show the new intervention is better than nothing [13,102].

“Often, NI trials in and of themselves do not address unmet medical needs of study patients because, by definition, the existing standard intervention is already known to be effective and thereby is addressing the need for those patients in the study.”

Even if measures are taken to avoid pitfalls in design, some authors have held that NI trials are inherently unethical when patients are asked to participate in trials that will not provide any advantage to them [17]. While US regulations indicate that consent forms should inform research subjects of the purpose of the research and other available alternative therapies [101], few consent forms inform subjects that NI trials evaluate how much worse a new intervention might be than Standard. Would patients enroll in NI trials if they understood they could be randomized to an intervention that is 10–15% less effective than interventions they could already receive? Furthermore, federal regulations also spell out that Institutional Review Boards should evaluate risks and benefits resulting from the proposed research, and not simply consider possible long-range effects of applying knowledge from the research. For example, this calls into question the notion of conducting NI trials of experimental antibiotics in susceptible populations based on interests in identifying options for current or future patients with resistant...
Noninferiority trials clinical understandings & misunderstandings

Conclusions & future research
These considerations motivate the importance of better education and information for patients, investigators, clinicians, regulators and Institutional Review Boards regarding the goals of NI trials and the types of research questions they should and should not be used to address. It is important that there be an understanding of the attendant risks for research subjects when studying a new intervention that is not hypothesized to have better efficacy than the established effective Standard, and yet with considerable likelihood could be meaningfully less effective. As part of an evidence-based formulation of the NI margin, future research is needed on patients’ views regarding how much loss of effect with a new intervention is clinically acceptable under different settings. Then, to ensure truly informed consent, the consent forms for trials should clearly acknowledge the research goal for the NI trial is to distinguish between hypotheses that the new intervention is less effective than Standard (at the level specified by the NI margin) versus being equally effective. Finally, there is a need to recognize trial designs other than NI should be used whenever possible, and certainly when there is an absence of reliable historical evidence regarding the effect of Standard in the setting of the NI trial.

Financial & competing interests disclosure
TR Fleming, received financial support for research described in this article from a National Institute of Health/National Institute of Allergy and Infectious Diseases grant entitled “Statistical Issues in AIDS Research” (R37 AI 29168).

JH Powers is funded in whole or in part by federal funds from the National Cancer Institute, National Institutes of Health, under contract N01-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. This research was supported in part by the National Institute of Allergy and Infectious Disease. Collaborative Clinical Research Branch, JH Powers, MD, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References
10 Davis KS. Non-constancy, estimation bias, biocreep, and an alternative to current methods used in non-inferiority trials. University of Washington, WA, USA (2010).
17 Garratini S, Bertele V. Non-inferiority trials are unethical because they disregard patients’ interests. Lancet 370, 1875–1877 (2007).