ELSEVIER

Journal of Clinical Epidemiology 66 (2013) 386−396

Journal of Clinical Epidemiology

About half of the noninferiority trials tested superior treatments: a trial-register based study

Beryl Primrose Gladstone*, Werner Vach

Clinical Epidemiology Group, Department of Medical Biometry and Statistics, University Medical Center Freiburg Stefan-Meier-Str. 26, D-79104 Freiburg, Germany

Accepted 24 October 2012; Published online 18 January 2013

Abstract

Objectives: A concern that noninferiority (NI) trials pose a risk of degradation of the treatment effects is prevalent. Thus, we aimed to determine the fraction of positive true effects (superiority rate) and the average true effect of current NI trials based on data from registered NI trials.

Study Design and Setting: All NI trials carried out between 2000 and 2007 analyzing the NI of efficacy as the primary objective and registered in one of the two major clinical trials registers were studied. Having retrieved results from these trials, random effects modeling of the effect estimates was performed to determine the distribution of true effects.

Results: Effect estimates were available for 79 of 99 eligible trials identified. For trials with binary outcome, we estimated a superiority rate of 49% (95% confidence interval = 27−70%) and a mean true log odds ratio of −0.005 (−0.112, 0.102). For trials with continuous outcome, the superiority rate was 58% (41−74%) and the mean true effect as Cohen's *d* of 0.06 (−0.064, 0.192).

Conclusions: The unanticipated finding of a positive average true effect and superiority of the new treatment in most NI trials suggest that the current practice of choosing NI designs in clinical trials makes degradation on average unlikely. However, the distribution of true treatment effects demonstrates that, in some NI trials, the new treatment is distinctly inferior.

Keywords: Noninferiority trials; True treatment effect; Inferiority; Effect estimate; Biocreep; Risk of degradation

1. Introduction

Noninferiority (NI) trials attempt to demonstrate that the loss in efficacy by a new treatment compared with a standard treatment is limited to a certain prespecified margin. The new treatments being tested usually possess some form of advantage over the standard one, which allows a tradeoff against a small drop in efficacy. The presumed advantage may be in the form of safety, ease in administration, tolerability, costs, and so on [1−3]. There has been a surge in the use of NI designs in the past decade [4], and, moreover, evidence based on NI trials is being used for new drug approval [5], all of which subsequently affects the availability of drugs in the mainstream and clinical practices thereafter.

A matter of concern is that NI trials impose a risk to accept a new treatment, which is not superior to the standard treatment or in other words, having a negative true treatment effect. In contrast to superiority trials, this risk is not limited by the significance level chosen. Since the early days of NI trials, a concern about the risk has been expressed by researchers, often referred to as biocreep [1−3,6−10]. Biocreep basically refers to the cyclical phenomenon where a slightly inferior treatment becomes the active control for the next generation of NI trials, which over time leads to degradation of the efficacy of the treatment offered to patients [1]. Although it has been argued that, theoretically, this cyclicality toward biocreep cannot be maintained and thus is not plausible [9]; NI trials still pose a risk of degradation, on an average, of the treatment success in areas where NI trials are popular.

Although this risk depends on many aspects, including the choice of NI margin, assay sensitivity, adherence of constancy assumption, quality of trial conduct, and so on, the foremost contributor is the distribution of true effects in NI trials. The distribution of true effects would reveal how often the current NI trial study design is being chosen

Conflicts of Interest: The authors declare that they did not get any support from any organization for the submitted work and did not have any financial relationships with any organizations that might have an interest in the submitted work.

* Corresponding author. Tel.: +49-761-203-6658; fax: +49-761-203-6711.

E-mail address: beryl@imbi.uni-freiburg.de (B.P. Gladstone).

0895-4356/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jclinepi.2012.10.011

What is new?

- Noninferiority (NI) trials impose a risk of accepting a new treatment, which is not superior to the standard treatment, and researchers have expressed their fear that this risk may lead to degradation of the efficacy of available treatments.
- One of the major contributors for this risk is the distribution of true treatment effects in NI trials. Our result that the average true treatment effect is positive suggests that the current practice of choosing NI designs in clinical trials makes degradation on average unlikely.
- Our finding that the rate of testing superior treatments among all NI trials is approximately 50% suggests that trialists do not systematically compare less effective new treatments with standard treatments.
- Our finding suggests that NI trials are equally done with superior treatments with a small treatment effect and non-superior treatments.
- However, the estimated 95% ranges suggest that in some NI trials, new treatments are tested, which are distinctly inferior to the standard treatments.

to compare less effective new treatments with standard treatments. Generally speaking, it would reflect partly the current practices of overall clinical research. The danger of biocreep expressed in recent years [2,3,6−10] implies that some scientists feel that true effects in most NI trials are negative.

In this context, Soonawala et al. [6] studied a set of 170 published NI trials aimed at assessing the hypothesis that new treatments that gain a verdict of NI are systematically less effective than standard treatments. They observed an average estimated treatment effect close to zero contradicting their hypothesis. The finding implies that the current practice of choosing NI designs in clinical trials makes degradation on average unlikely.

However, the authors themselves mention the possibility of publication bias, and that the inspection of effect estimates in published NI trials could be misleading. Public registries of clinical trials and their results [11] open possibility to study an unselected set of NI trials and hence lowering the probability of publication bias.

Hence, to supplement the finding of Soonawala et al. [6], we aim to study the effect estimates from a set of all registered NI trials conducted within a 7-year period to determine the distribution of true treatment effects. We present the estimated fraction of NI trials actually investigating superior treatments among all trials designed to be an NI trial and also present various aspects of the distribution of true effects.

2. Methods

The research work comprised identification of NI trials from a clinical trials register followed by search for results from these trials with the help of various sources, extraction of effect estimates from the result sources, and the analysis of effect estimates to determine the distribution of true treatment effects. Figure 1 summarizes this overall research strategy.

2.1. Identification of NI trials

We identified all NI trials carried out (started and completed) between January 2000 and December 2007 among the trials registered either in the National Library of Medicine (NLM) clinical trials register (www.ClinicalTrials.gov) developed by the US National Institutes of Health [12] or the *International Standard Randomized Controlled Trial Number (*ISRCTN) register maintained by the Current Controlled Trials, Ltd. [13]. The search was carried out in all the sections of the register using the terms ‘‘non-inferiority,’’ ‘‘noninferiority,’’ ‘‘non inferiority,’’ and ‘‘not inferior’’ in advanced search option of the register. The period was restricted until December 2007 so that all studies had a reasonable chance to have published the results. Studies conducted before 2000 were excluded to ensure that we investigate the current practice as much as possible. The exclusion of trials outside 2000−2007 was done using the search option at the NLM register. But for the ISRCTN register, all the trial protocols were screened manually as the search option does not include trial start and completion dates. Data from the registered trial protocols were extracted (details in Appendix at www.jclinepi.com).

2.2. Inclusion and exclusion criteria

All the identified trials were carefully screened to include only clinical trials studying NI of efficacy of a new drug/treatment/therapeutic procedure/diagnostic procedure as the primary objective. The NI trials aimed at determining the optimal dose of a drug with no comparison with a standard drug were excluded. Vaccine trials were excluded post hoc because they typically have many primary endpoints studying various strain-/subtype-specific antibodies and often consider protective rates close to 100%, making the estimation of effect and standard error unreliable.

2.3. Search for study results

We looked for the trial results in various sources, which included the trial registers’ results section; www.