Noninferiority is almost certain with lenient noninferiority margins

Dear Editor,

We thank Dr. Soonawala et al. for their interesting comment [1] to our article on the sponsorship of head-to-head randomized trials [2]. Our proportion of noninferiority trials with “desirable favorable results” (97%) was higher than the one reported by Soonawala (84%, if we only consider the 165 trials with a sample larger than 100, that was one of our study selection criteria) [2,3].

The observed discrepancy may still be due to chance, or it may be caused by other differences between the two studies. The previous analysis [2] was based on trials published between 1999 and 2009, and we examined trials published in 2011. It is possible that the machinery of the industry producing favorable results from noninferiority trials may have become more efficient recently. Moreover, Soonawala et al. analyzed only trials published on core journals, whereas 45/57 of our trials were not published on core journals, which increased the likelihood of including some more biased trials. Of the 12 trials published in core journals, 11 (92%) had favorable results, not significantly different from the proportion in Soonawala et al. [2].

Overall, we fully agree with the concluding statements of Soonawala et al. [1]: the noninferiority design seems a very “safe” design, and the rationale for choosing it should be more consistently reported in publications and registries, as it was also more recently documented [4]. In addition, we share previous concerns on the choice of noninferiority margins [5,6]. Among the 33 comparisons in which the result was expressed as a relative risk in Soonawala et al. survey, 6 choose a margin equal or higher that 2.00 [3]. Although the selection of the noninferiority margin is based on clinical judgment in addition to statistical reasoning, these trials selected margins that were too lenient. To provide the highest standard of patient care, researchers must be very careful in designing noninferiority trials.

Maria Elena Flacco
Department of Medicine and Aging Sciences
University of Chieti
Via dei Vestini 5
66013 Chieti, Italy

Regional Health Care Agency of Abruzzo
Via Attilio Monti 9
Pescara, Italy

Lamberto Manzoli
Department of Medicine and Aging Sciences
University of Chieti
Via dei Vestini 5
66013 Chieti, Italy

References

http://dx.doi.org/10.1016/j.jclinepi.2015.11.010

Noninferiority is (too) common in noninferiority trials

In their meta-analysis of 317 randomized trials that examined efficacy and safety of drugs, biologics, and medical devices in head-to-head comparisons, Flacco et al. [1] report that 97% (55 of 57) of industry-funded noninferiority/equivalence trials provided favorable results (defined as superiority or noninferiority of the experimental treatment) compared to 71% (20 of 28) of nonindustry-funded noninferiority/equivalence trials. We reviewed the data from our previously published meta-analysis of 175 noninferiority comparisons [2]. Overall, 83% (145 of 175) of the comparisons produced desirable favorable results (superiority or noninferiority for the experimental treatment). In contrast to Flacco et al. [1], there was no important
difference in this percentage by type of funding source. Fig. 1 shows the combined risk ratios from random-effects meta-analyses by funding source, and the percentage of comparisons with favorable results. There was little evidence for heterogeneity in combined risk ratios ($P = 0.20$ from chi-squared test) and percentages ($P = 0.49$ from chi-squared test). The trials examined by Flacco et al. [1] were published in 2011, whereas our analysis was based on trials published between 1993 and 2009. It is therefore possible that the difference between industry-sponsored and other trials emerged in recent years. However, in our data set, we found no clear trend over time in this difference.

Taken together, the data from these two meta-analyses show that a large majority of published noninferiority trials support the experimental treatment. The probability of getting a verdict of noninferiority in a noninferiority trial is greater than 80%. From the viewpoint of industry and other stakeholders, the noninferiority design is therefore a very “safe” design. The situation is more variable for superiority trials. The study by Flacco et al. [1] and another meta-epidemiological study of trials from a range of medical specialties found that experimental treatments were superior in 68% to 74% of trials [3]. In contrast, two studies in oncology found that only 25% to 45% of trials produced results compatible with superiority of the experimental treatment [4,5].

There are valid reasons for choosing a noninferiority design when performing a trial. However, industry and other stakeholders may be tempted to choose the design for the wrong reasons; that is, to increase the chance of obtaining a favorable result even when a superiority design would have been more appropriate to answer the questions that really matter to patients. Therefore, the rationale for choosing the noninferiority design, the primary outcome, the noninferiority margin and the evidence for the supposed benefit of the experimental treatment should always be specified and consistently reported in study protocols, trial registries, and publications [6–8].
Effects of general health checks differ under two different analyses perspectives—the Inter99 randomized study

Recently, a large randomized intervention—the Inter99—has confirmed that health check followed by repeated lifestyle counseling has no effect on 10-year mortality at a population level [2]. Despite this, many researchers, health professionals, and politicians still believe that preventive health checks are effective in preventing disease and death in the general population [3]. General preventive health checks followed by repeated counseling of high-risk persons are currently implemented at a national level in several countries (e.g., England, Japan, Russia, and Australia). The evidence that favor an implementation of preventive health checks primarily comes from quasi-experimental studies which describe changes in, for example, lifestyle and self-reported health or blood pressure and not in hard end points such as death [4-7]. Furthermore, these conclusions regarding a beneficial effect have primarily been based on analyses restricted to participants [8,9]. However, excluding nonparticipants after randomization may introduce noncomparability of the control group (CG) and intervention group (IG) and lead to biased results. Recently, an article has been published showing that selection bias in large part influences the effect measures of the Inter99 study [10].

The Inter99 study, a population-based, randomized lifestyle intervention, took place in the years 1999-2006 (for a detailed description please see [11]). The study population (age 30-60 years) comprised 11,629 persons in the IG and 47,987 persons in the CG. At baseline, a random sample of the control group (SCG) (n = 5,228) received questionnaires about health and lifestyle. The remaining CG was not aware of taking part in an intervention study. Persons defined as being at high risk of ischemic heart disease were repeatedly offered individual and group-based lifestyle counseling over a period of 5 years. All persons

Funding: This work was supported by the Health Foundation (Helsefon- den) Fondenes Hus, Otto Mønsteds Gade 5, 1571 København V, Danmark (grant 2010 B 131).

Conflict of interest: None.