Author's response to reviews

Title: Phase IV non-inferiority trials and additional claims of benefit

Authors:

Rosemarie DLC Bernabe (R.Bernabe@umcutrecht.nl)
Grace Wangge (g.wangge@uu.nl)
Mirjam J Knol (mirjam.knol@rivm.nl)
Olaf H Klungel (O.H.Klungel@uu.nl)
Johannes JM van Delden (j.j.m.vandelden@umcutrecht.nl)
Anthonius de Boer (a.deboer@uu.nl)
Arno W Hoes (A.W.Hoes@umcutrecht.nl)
Jan AM Raaijmakers (jan.a.raaijmakers@gsk.com)
Ghislaine JMW van Thiel (g.j.m.w.vanthiel@umcutrecht.nl)

Version: 4 Date: 7 February 2013

Author's response to reviews: see over
Dear Editor,

Thank you for allowing us to revise our manuscript. I think the reviewers’ comments have helped a lot to make our manuscript better. Below you will find our responses in blue.

Authors’ reply 1012486587804563- R1 - Phase IV non-inferiority trials and additional claims of benefit

Reviewer 1: Jörg Zinserling

Major comments/ Major Compulsory Revisions

1. The authors present a manuscript on an analysis of non-inferiority (NI) trials which can be considered Phase IV trials in a broader sense. The research question is how “additional benefit” is addressed in the trials and on which basis “additional claims” are justified. The research question is relevant.

2. The sample of studies included in the author’s analysis is originating from the selection of studies used for a previous review of the group. The selection procedure is comprehensively described and the selection process leading to the selection of 15 definite Phase IV and 35 post-authorization studies is adequate. Further classification into pharmaceutical industry sponsored and non-industry sponsored studies is helpful. The authors adequately address the limitation of the sample size in the discussion. Some aspects of the terminology and statements by the authors need further explanation.

We thank the reviewer for his thorough review and helpful comments.

3. It is not clear to me how the authors define the term “formal test” as used in the results section (p. 5). In the context of inferential statistics in my understanding a formal (statistical) test could be a test with pre-defined hypotheses that are stated in advance and evaluated with assessment of p values size and precision of the treatment effect and its clinical relevance. This would be a formal test needed in an adequately controlled trial with confirmatory objective and documented sample size estimation (cf. ICH E9 on Trial Context). A formal statistical test could also meant be used in an exploratory sense where choice of hypotheses may be at least in part data dependent like in trials which use not only descriptive statistics for the secondary endpoints. The authors need to clarify what they mean by “formal test”.

We agree with the reviewer that the term “formal test” can be confusing and we now explain more explicitly what we mean in the revised version of the paper. In essence we mean that there is an a priori objective/hypothesis pertaining to the additional claim, preferably a sample size calculation and data analysis plan. Further, and following the lead of ICH E9, we used “statistical analysis” instead merely of “test” since the term “statistical analysis” was used throughout that document.
4. I would not agree with the authors that Phase IV NI trials (with assay sensitivity secured) need in principle to have additional benefit claims that are to be based on confirmatory “formal tests”. Trials with the primary objective to show NI to a comparator can contribute to the “improvement of therapeutic interventions” (p. 7) also by exploratory analyses based on statistical tests which would not be expected to be confirmatory in nature (see comment 2). Furthermore, assessment of the safety profile is usually not based on pre-specified hypotheses with claims stated in advance. The authors need to clarify in which sense the term “additional claim” is meant.

We agree with the reviewer that “Trials with the primary objective to show NI to a comparator can contribute to the ‘improvement of therapeutic interventions’ also by exploratory analyses based on statistical tests which would not be expected to be confirmatory in nature.” However, for additional benefit claims in Phase IV NI trials to be clinically relevant and informative for formal decision-making by regulatory and reimbursement agencies, these claims must be confirmed using confirmatory statistical analysis. Precisely because it is a claim, it makes much sense to think that there were previous exploratory studies before such claims could be made; and second, that such claims must be tested in trials aimed at confirming (or refuting) the claims.

To avoid misunderstanding, in the introduction, instead of saying, “As such, in principle, all NI trials should have additional benefit claims”, we now say, “As such, in principle, all NI trials should aim at specifying and demonstrating additional benefit claims.”

By additional benefit claims, we refer to any other claim aside from the (usual) efficacy endpoint. It may be safety, ease of use, etcetera. We believe our statement in the introduction (which reads, “Such additional claims may relate to improved safety, but also optimization of the method of administration, improved compliance, and cost-effectiveness.”) captures what we mean by additional benefit claims.

5. The authors limited their analysis of trials to NI Phase IV trials. If “additional claims” would be meant by the authors to be based on formal tests in a confirmatory sense a trial with this objective could just as well be a trial aimed to show superiority as a primary objective. By definition, superiority Phase IV trials are not included in the presented analysis. The authors should comment on this in the context of answers to the comments 3. and 4. and possibly include this aspect in the discussion.

We agree with the reviewer’s comment. Indeed, we wish to direct the reviewer’s attention to an article in Drug Discovery Today that is authored by some of us entitled, “Should non-inferiority trials be banned altogether?” (Please see Wangge G, Klungel OH, Roes KC, de Boer A, Hoes AW, Knol MJ.
Should non-inferiority drug trials be banned altogether? Drug Discov Today. 2013 Jan 14. pii: S1359-6446(13)00005-6. doi: 10.1016/j.drudis.2013.01.003). This article suggests that NI trials may be designed in a manner that simultaneously shows the NI objective “with regard to drug efficacy and the objective of establishing superiority of the additional advantages of a drug over its active comparator.” This is the context of our remarks. To properly position our paper, we added the following in the discussion section:

“In addition, in a previous article, we pointed out that NI trials may be designed in a manner that simultaneously shows the NI objective ‘with regard to drug efficacy and the objective of establishing superiority of the additional advantages of a drug over its active comparator.’ This may be a viable direction to follow if we are to revise the methodology of phase IV NI trials.”

Minor Comments/ Discretionary Revisions

6. The authors should comment on the possibility to include phase IV trials that are not explicitly non-inferiority trials in their review.

We performed a search in PUBMED using the search terms, “non-inferior*”, “noninferior*” or “active control and “equivalence”, in combination with the MeSH term “humans” and “Randomized Controlled Trial” as publication type. By doing that, we aimed to include only articles on non-inferiority trials in our analysis. The phases of the trials were later extracted from the selected articles. Thus, we believe our analysis was free from the possibility of including phase IV trials that are not explicitly NI trials.

7. The message that Phase IV trials should be improved in methodology is supported by this reviewer.

We thank the reviewer for this positive comment.

Reviewer 2: Gerd K Rosenkranz

Minor essential revisions:

8. The authors correctly state that any NI trial should be able to demonstrate some additional benefit of the experimental treatment over some reference treatment above and beyond being not less efficacious than the reference in terms of a primary efficacy endpoint; otherwise the new treatment would mainly provide a contingency in case a patient cannot tolerate the reference treatment. Besides the important aspect of clinical relevance, aiming for demonstrating a benefit of the new treatment on endpoints other than the primary can help to relieve some of the criticism raised against NI trials like no direct evidence of assay sensitivity (see Temple and Ellenberg, 2000).

9. The authors mention that they randomly selected 300 of 669 articles “based on pragmatic consideration rather than formal sample size calculations”. Was 300 the amount of articles that could be processed within some time frame? What are the implications of this selection for the results of the paper?
We thank the reviewer for his thorough review and helpful comments.

As the reviewer suggested, the decision to study a random sample of 300 papers (and not to scrutinize all 669) had to do with the workload consideration and the fact that 300 manuscripts would suffice to arrive at valid conclusions. We think we have clarified this in the methods section when we said, “This search resulted in 669 articles and, based on pragmatic consideration rather than proper sample size calculations, we randomly selected 300 for our review.”

As stated in the manuscript, the data used is the same as that in a previously published manuscript (please see Wangge G, Klungel OH, Roes KC, de Boer A, Hoes AW, Knol MJ. Interpretation and inference in noninferiority randomized controlled trials in drug research. Clin Pharmacol Ther 2010 September;88(3):420-3). Hence, our current analysis was part of previous analysis done in 2009, where we randomly selected 300 articles from 669 articles found in Pubmed on 5th February 2009. As expected, since the 5th of February 2009, there were more articles on NI trials that were published and indexed in Pubmed. Thus, we missed recently published Phase IV NI trials.

In the revised version, we added the information on the search date and addressed this issue as one of the limitations of our article. In the discussion section, the added limitation reads as follows:

“Another limitation will be the fact that our current analysis was part of a previous analysis where we randomly selected 300 articles from 669 articles found in Pubmed on the 5th February 2009. As expected, since that date, there were more articles on NI trials that were published and indexed in Pubmed. Thus, we missed recently published Phase IV NI trials. However, we do not believe the main message of our article, i.e., the importance of showing additional benefits in Phase IV NI trials, will differ much with the addition of those new articles in our analysis.”

10. The authors classify studies as pharmaceutical industry sponsored or non-pharmaceutical industry sponsored according to whether industry was actively involved in writing the protocol, monitoring, analysis or reporting or not. In case a trial was funded by industry without further involvement, it was also classified as non-industry sponsored. This is considered misleading since funding constitutes one way of sponsoring. For the sake of clarity the authors should consider a classification according to industry-initiated (including funding and further involvement), non-industry initiated but industry funded, and non-industry sponsored trials.

We agree with the reviewer that categorizing the type of sponsor could be misleading, thus we classified the studies based on the type of the trial’s initiator, as also suggested by this reviewer. In our analysis, we classified the studies in pharmaceutical-industry-initiated or non-pharmaceutical-industry-initiated studies to prevent any confusion. This change is reflected in Table 3.

11. It is recommended to avoid the terminology of a “formal” statistical analysis throughout since there might be nothing like an “informal” analysis. The terms “descriptive” versus “confirmatory” should be more appropriate. Likewise the authors should be mindful about the meanings of “efficacy” and “effectiveness”.

Regarding the term “formal”, we do not use it in a relative manner, as in the opposite of “informal”. Rather, we use the term in the following sense as stated in the online Merriam-Webster Dictionary: “following or according with established form, custom, or rule.” Given our context, this would mean
following or using established statistical testing/analysis. Please also see our response on comment #3 since instead of using “formal test” we now use “formal statistical analysis”.

Regarding the terms “efficacy” and “effectiveness,” we checked the whole article for the terms “efficacy” and “effectiveness” and made necessary revisions.

12. In 73% of the trials was non-inferiority concluded. Was additional benefit only investigated in these trials or also in those where non-inferiority could not be confirmed? It would also be interesting to know what analysis population was used for the NI analysis and whether the NI margin was reasonably justified.

Additional benefit was investigated on all trials that were categorized as phase IV NI trials and not only in the 73% where non-inferiority was concluded.

We added the additional benefit data and the conclusion of the trials in table 3 and the type of analysis that were used in table 1. We do not see a clear pattern that the conclusion of the trials was different across the types of additional benefits claimed. Thus, we did not discuss this further.

13. Only one trial performed a sample size calculation for the claimed additional benefit. Rosenkranz (1997) is another industry-initiated one.

We checked and found that Rozenkranz (1997) was indeed not included in our analysis. It was not included in our analysis because the trial was a dose-finding trial and thus categorized as a phase III trial.

14. The authors mention the concept of “comparative effectiveness” which is becoming more relevant in the assessment of new treatments. This approach is complemented by the introduction of patient reported outcomes (PROs). PROs let the patients speak about what they consider a benefit as compared to what the caregivers consider and provide a tool to measure additional benefit. The inclusion of PROs early on in the development process is strongly supported by health authorities, for example by the Food and Drug Administration.

We agreed with the reviewer that PROs could be one of the important tools to measure the additional benefit. However, we limited the scope of our article to describe how post-authorization NI trials reported benefit claims beyond clinical efficacy and how these additional claims were supported or proven in the trials. We think that the discussion on which means may be used to show additional benefit claims will need further analysis and may best be done in another paper.

15. The terminology “establish clinically relevant differences” is ambiguous. It could mean to demonstrate a superior effect in an endpoint other than the one in which NI was established. It could also mean provide evidence for a relevant difference in addition to a statistically significant result which provides evidence for some difference. The authors are asked for checking the document to achieve clarity where applicable.

By clinically relevant differences, we refer to differences that are potentially of “clinical significance” (please see ICH E9 section 2.1.2), i.e., the first sense mentioned by the reviewer. It means confirming the additional benefit claim. Please also see our response to #5.
16. The authors raise ethical concerns about trials that only demonstrate non-inferiority in regard to existing treatment without the intention to demonstrate additional benefits. They should consider the following in this context. First, a treatment which is not worse than an existing one is doing no harm. At least in indications with a medical need and a small number of effective treatments, an additional treatment option may still be worthwhile even if it cannot be proven to be better than existing ones in some important variables. It provides the treating physician with an alternative in case existing treatment is not tolerated by an individual patient. Second, in many situations combinations of drugs with similar efficacy may prove more efficient in combination but have to prove efficacy on their own. Third, newly developed treatments are usually more expensive than existing ones (in particular when the latter have become generic). The prospect that payers may not be willing to pay additional money for the same value is a substantial hurdle to develop new medicines without consideration (and confirmation) of additional benefit.

Regarding the reviewer’s first point, we agree that having alternatives is indeed useful; however, there is still the need for additional benefit claims because of the following:

A. Even among me-too drugs, the claims of one drug being more tolerable for individuals with specific conditions ought to be established.

B. Showing non-inferiority for only one endpoint does not prove that the drug is not worse than the comparator. This might be but it is also possible that the new drug is worse than the comparator for not evaluated outcomes.

C. The trial may be burdensome or even have risks (extra diagnostic tests etc.) This is a relevant consideration for the justification of phase IV trials.

Regarding the second point, this reviewer may be right that the combination of drugs with similar efficacy may be more efficient, but that we think is beside the point we are making in this manuscript.

As to the third point, we think the reviewer is arguing for our favor since we agree that it is a hurdle for pharmaceutical industry to develop new drugs with no additional benefit claims, especially if these new drugs would be more expensive. However, as we can see from our data, me-too drugs are continuously developed with additional benefit claims that are hardly confirmed. Hence, there is the need to talk about it in reputable journals such as BMC Medical Research Methodology.


Comments from the Editor

18. The arbitrary selection of 300 papers is a weakness of this piece of research. Given that the results section is with less than one page fairly short the authors might consider extending their review. Furthermore, the reviewers make a number of excellent suggestions that need to be addressed in a revision.

Please see comment #9.

Additional Editorial Requirements:
19. Tables as figure files: You have uploaded the tables as figure files. Please remove them from the submission system and include the tables within the text file of the manuscript after the references. The tables should be formatted using the Table tool in your word processor. Please also move the table title to above the table and the legend to below the table, within the text.

Corrected.

I hope our revisions and responses are sufficient for our manuscript to be accepted by BMC Medical Research Methodology.

Thank you!

For the co-authors,
Kind regards,

Rosemarie Bernabe, MA