Author's response to reviews

Title: Calculation of NNTs in RCTs with time-to-event outcomes: A literature review

Authors:

Mandy Hildebrandt (Mandy.Hildebrandt@iqwig.de)
Elke Vervoelgyi (Elke.Vervoelgyi@iqwig.de)
Ralf Bender (Ralf.Bender@iqwig.de)

Version: 3 Date: 29 July 2008

Author's response to reviews: see over
Response to the Reviewers

28.07.2008

Response to Reviewer 1 (Jacobus Lubsen):

1) We have not overlooked that NNTs must be defined for a certain duration of follow-up. A complete and correct definition of NNT was given on page 6 of the previous version of the manuscript: "... in this paper the NNT is – as usual – considered as effect measure comparing the risks of two groups for a specific amount of follow-up time (which is not necessarily identical with the duration of treatment) in terms of patient numbers having to be treated to expect an avoided event in one patient."

Nevertheless, to avoid any misunderstandings we added the dependence of NNTs on follow-up time in ALL descriptions of NNT (pages 4 and 5). We also deleted the between-brackets addition "regardless of treatment time" on page 5 to avoid misunderstandings.

2) Our statement in the first response that the reviewer equates by mistake follow-up time with treatment duration was NOT based upon the Lancet 2000 paper. It was based upon the reviewer's first report. He wrote "... The reason is that patients who die are treated for less than three years even though the follow-up was the same for all patients in the trial concerned. Hence, despite the fact that the follow-up was the same for all patients the NNT has to be derived from the underlying hazards. Assuming exponential survival ...". In these statements the reviewer obviously equates the observed person-time with treatment duration. Unfortunately, in our statement we used the term "follow-up time" rather than observed person-time, which was imprecise because it can be confused with study time. Thus, a better formulation of our statement is "The reviewer equates by mistake observed person-time with treatment duration." It is correct that Lubsen and coworkers explained in the Lancet 2000 paper that person-time is not necessarily identical with treatment duration. Thus, the use of hazards does not automatically lead to an effect measure that can be applied for cost-effectiveness analyses. However, this is not the main point.

The main question is whether in the case of chronic diseases and continuous treatment the calculation of NNTs HAS TO BE based upon hazards. The argument of Lubsen and coworkers is that an effect measure is required which is adequate for cost comparisons. We do NOT follow this argument. NNT is an absolute effect measure to quantify the effect of a treatment in terms of patient numbers. Sometimes this measure is adequate for cost-effectiveness analyses and sometimes not. This depends on the circumstances. It is misleading to define an alternative effect measure (in terms of person-time rather than person numbers) and use the same name. If the NNT is inadequate for a specific purpose such as cost-effectiveness analyses one should use an alternative effect measure but one should also use an alternative term to avoid misunderstandings.
We revised the corresponding statements in the background of the manuscript, because they were misplaced due to our attempt to deal with the first report of reviewer 1 (see also point (i) of reviewer 2). We moved the discussion of the treatment duration issue to the discussion section.

General comments:

3) The statement that “only in the case of exponential survival times and rare events the hazard difference can be used to approximate risk difference” is not nonsense, it is a fact (with the addition that - in theory - also the linear hazard rate distribution is possible, see point 6). This is based upon the same logic with which one can use odds ratios estimated from case-control studies as approximations of relative risks. It may be that it was not the reviewer's goal to approximate the NNT by the inverse of the hazard difference. However, our general statement remains correct.

It is trivial that one can calculate risk differences from hazard differences if the information needed to do so is available. However, the information needed is the correct survival time distribution. The knowledge of only the hazard difference for one specific time point is insufficient to calculate the corresponding risk difference.

As suggested by the reviewer, we used the provided table with numerical examples to think our points through and found that the table actually confirms our statements: (i) NNT is dependent on time, (ii) for constant hazards and low risks the inverse of the hazard difference approximates the NNT.

4) The main objective of the paper is to investigate whether the methods used for NNT calculation in RCTs with time-to-event outcomes published in medical journals are appropriate. Thus, it is inadequate to "list the various NNT estimation methods that have been used in the literature in neutral manner". To investigate the frequency of inappropriate calculation methods we have to explain which methods are appropriate and which are not in the methods section.

5) The reviewer likes to see a number of additions to the supplemental table and suggests that this table be part of the manuscript. Most of the proposed additional columns represent additional study details which are beyond the scope of our paper. We added one new column with the information about confidence intervals for NNTs. This is a useful suggestion because it is in line with our goals. We are willing to include the table as part of the manuscript if the Editor thinks this is appropriate. However, we think that this table containing details of the study characteristics is better placed as additional file.

6) The terms "exponential distribution" and "constant hazard" are equivalent because the exponential distribution is the only distribution with constant hazard function. In principle, it is correct that it would be sufficient that the hazard difference is constant rather than the hazards of both groups. Thus, we added that also the linear hazard rate distribution (which could lead to constant hazard differences although the hazard itself is time-dependent) is a valid distribution (page 5). However, this is a purely theoretical point without practical importance, because the exponential distribution is the only survival time distribution leading to a constant hazard difference that is commonly used in medical applications. Notably, ALL calculations of the reviewer in the reports and the Appendices of the Lancet 2000 paper are based upon the exponential distribution.
7) It is correct that we do not know definitively that, e.g., the authors of the EUROPA report used an appropriate calculation method for the NNT. In the case that the reported NNT equals the recalculated NNT from the reported survival probabilities we classified the method as "appropriate" although it may be that in fact naive proportions have been used but the incorrectly calculated NNT is haphazardly identical to the correct NNT. We added this point to the methods section. We also added the fact that this may lead to an underestimation of the proportion of papers with inappropriate NNT calculations to the discussion.

Manuscript:

8) changed
9) changed
10) changed
11) changed

12) We added that the calculation of NNTs based upon survival probabilities estimated by the Cox model as described by Altman & Andersen is also appropriate. However, we are unable to distinguish between the two estimation methods proposed by Altman & Andersen, because only one article described in detail which of the two methods has been used.

13) Parts of page 5 were moved to the discussion.

14) We added the information about confidence intervals to the supplemental table. We also added the information whether confidence intervals of the risk difference are provided if confidence intervals for NNTs are not given.

15) Based upon our definition (which is the usual one) NNTs are only NNTs when the unit is "number of patients". The study quoted on page 7 neither used the term "NNT" nor gave a result in terms of patient numbers. Consequently, this study was classified as non-NNT-reporting article.

16) We mentioned Table 3 in the methods section. We did not identify the reports in this table because the message of the table is not who made something wrong. The aim of the table is to provide information about the distribution of the differences between reported and recalculated NNTs.

Response to Reviewer 2 (Jorgen Nexoe):

Major Compulsory Revision:
1) We added a statement about the usefulness of NNTs to the conclusion.

Minor Compulsory Revision:
2) We revised the background.
3) changed
4) changed
5) We performed a thorough proof-reading.