Reviewer's report

Title: Use of methods to calculate the number needed to treat in randomised controlled trials where the outcome is time to an event: literature survey

Version: 1 Date: 11 February 2008

Reviewer: Jacobus Lubsen

Reviewer's report:

Major compulsory revision (one long comment):

The authors must be commended for the effort they made to collect and rate 734 trial reports. This is an interesting manuscript that eventually deserves to be published. My problem is that the authors don’t define exactly what they mean by a NNT and what the purpose is of calculating one. Is it a measure of efficacy or a basis for comparing the cost-effectiveness of treatments? Because of this, it is not possible to judge the results of their survey of the literature.

Concerning the meaning of a NNT, the assertion in the introduction that “in randomised controlled trials with a binary outcome (where the follow-up time is the same for all patients) the calculation of NNTs is based on simple proportions (rates from a 2x2 table)” suggests incorrectly that meaningful NNTs can always be based on simple proportions when the follow-up is the same for all patients.

In this regard, a distinction must be made between a trial of an acute condition (e.g. acute myocardial infarction (AMI)) and a treatment of a fixed duration (e.g. thrombolysis) on the one hand, and trial of a chronic disease (e.g. patients who had an MI in the past) and a continuous treatment (e.g. a beta-blocker) on the other.

To illustrate, suppose that in a trial comparing thrombolysis with placebo in patients with AMI, the mortality after one month (the same for all patients) is 5/100 in patients assigned thrombolysis, and 10/100 in patients assigned placebo. The NNT is then 100/(10 − 5) or 20. This number is entirely understandable and can be directly related to cost: 20 courses of thrombolytic treatment are required to achieve that one additional patient survives the AMI.

Suppose conversely that in a trial comparing beta-blockade with placebo in patients who had an AMI in the past, the mortality after three years (the same for all patients!) is, as before, 5/100 in patients assigned beta-blockade, and 10/100 in patients assigned placebo. The same NNT of 100/(10 − 5) or 20 is now rather meaningless and it would certainly be wrong to say that “20 patients must be treated for three years to postpone one death”, as is customarily done. 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 with
constant hazard, these are \(\ln(0.95)/3\) or 1.71/100 patient-years for patients assigned beta-blockade and \(\ln(0.90)/3\) or 3.51/100 patient-years for patients assigned placebo, which corresponds to an NNT expressed in terms of patient-years of treatment required to postpone one death of \(100/(3.51 - 1.71)\) or 55.6 years. Had all patients been treated with assigned treatment until death or the end of follow-up, this can now also be directly related to cost: 55.6 years of taking beta-blockade. In reality, patients are rarely treated with assigned treatment until death or the end of follow-up. In the more usual case therefore NNTs based on hazards can only be translated to cost after correction for real use of treatment, as is explained in detail in the paper by Lubsen et al. that the authors quote.

In the introduction, the authors also assert that "the hazard difference is only a valid approximation of the risk difference if the survival times follow an exponential distribution and if event rates are low, for instance less than 5% [7, 8]." This is a again misunderstanding.

Two situations must be distinguished in this context. The first situation arises when the hazards are essentially constant over time. An NNT based on the hazard difference expressed in years is always correct in this case, and does not depend on the duration of follow-up. On the other hand, an NNT based on the risk difference is never correct for the simple reason that the risks, and therefore the risk difference, depend on the duration of follow-up even when the underlying hazards are constant over time (this is why Altmann et al. suggest in their 1999 paper to base NNTs on a "meaningful duration of follow-up", without saying how to come to a meaningful choice in this regard). Whether the "event rates are low" is completely irrelevant in this context and brings back the past, when epidemiologists used to talk about the "rare disease assumption". Because of this, it is inappropriate to consider NNTs calculated from Kaplan-Meier curves and NNTs calculated from the hazard difference as equivalent under certain conditions (as the authors do on page 6 in the methods section). Please note that NNTs calculated from Kaplan-Meier curves and NNTs calculated from the hazard difference have different dimensions, and can therefore never be equivalent!

The second situation occurs when the underlying hazards vary over time. In that case, the NNT will vary with time also, and will be different for, say, the first year of follow-up, the second year of follow-up, etc. (see the extra material that was on the Lancet’s website when the paper by Lubsen et al. was published).

In a revised version, the authors should first define what they mean by a NNT and what information a NNT is supposed to convey. They should then define explicitly which information is required to calculate meaningful NNTs, possibly distinguishing between trials of an acute condition and a fixed duration of treatment and follow-up, and trials of a chronic condition and a variable duration of treatment and follow-up. That being done it becomes a very useful exercise indeed to determine for how many of the 734 papers NNTs as defined by the authors were reported, and for how many these could be derived based on the information given by a reader who knows how to do that.
I wish the authors much success in revising this manuscript and I encourage them to do so, observing Voltaire’s suggestion "If you want to speak to me, define your terms".

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

No competing interests