Numbers needed to treat derived from meta-analysis

Are an absurdity

Using patient years may also be misleading
Length of follow up is poorly reported
NNT is a tool, to be used appropriately

Are an absurdity

EDITOR—Better late than never. Several years ago one group of epidemiologists put forward the number needed to treat (NNT) derived from megatrials and meta-analysis as a summary statistic suitable for expressing the effectiveness of medical interventions; now another group of epidemiologists has at long last realised that the NNT is seldom a valid measure.  

Some of us came to the conclusion that NNTs were "not necessarily true" rather more rapidly, and without the need for three and a half pages of cumbersome and dubiously appropriate statistical analysis. The deep flaws in the NNT statistic can be understood by a straightforward act of inference based on an understanding of the relevant clinical science and guided by the principle of "garbage in, garbage out."

The spurious precision of the NNT is a statistical artefact which derives not from clinical knowledge but from the illegitimate pooling of the large amounts of qualitatively unlike and clinically irrelevant data that are incorporated in almost all megatrials and meta-analyses. Unless trials incorporate patients with the same characteristics and the same prognosis and who are being given the same treatment as those to which the trial results will apply, then statistical summary is inevitably misleading.

It is somewhat galling that mega-epidemiologists and biostatisticians so routinely take credit for the act of creating spurious analytic tools, and then for belatedly dismantling them—but so it goes. The wheels of epidemiology grind exceedingly slow. At least Smeeth et al got there in the end.

When clinical epidemiology gives up its grandiose and self awarded claim to be "evidence-
based medicine” and once again becomes an activity based in clinical science, such absurdities may become a thing of the past. I hope so.

Bruce G Charlton, lecturer.
Department of Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU

Bruce.Charlton@newcastle.ac.uk


Using patient years may also be misleading

EDITOR—Smeeth et al showed how inappropriate methods of calculating numbers needed to treat (NNTs) in systematic meta-analyses can be misleading.¹ The examples they chose quoted event rates in patient years. Smeeth et al calculated (correctly in our view) their NNTs directly from the events per patient year. However, some commentators on such trials quote NNTs for the average follow up period of the trial. This alternative method may be considered acceptable, even though it is only an approximation to the first method, but we would draw attention to how misleading failure to recognise the difference between the two methods can be.

In the UKPDS 38 trial,² for example, method 1 would give an NNT to prevent any diabetic related death as 152 patients per year, or 15.2 patients over 10 years. Method 2 would give an NNT of 20 over 8.4 years (the median follow up). The three choices of NNT, 152, 15.2, or 20, can lead to misunderstanding. That this is a real problem was illustrated in an electronic response regarding the UKPDS 38 trial: “We are concerned that there is a discrepancy between the numbers needed to treat which are stated in the article, and those that can be calculated. The study states that the number needed to treat over 10 years to prevent any complication is 6.1 and to prevent death from a diabetes related cause is 15.0. In calculating the numbers needed to treat by using the values in figure 4 (based upon a median follow up of 8.4 years), we conclude that the number needed to treat to prevent any complication is 11, and to prevent death is 20.”³

At our local critical appraisal seminars for general practitioners in Suffolk, we encountered similar confusion when two participants presented NNTs from the HOT trial.⁴ An added twist to the potential for comparisons appears when some trials report the average follow up period as a mean (the HOT trial) while others report the median (UKPDS 38).
There is a case for standardising the way NNTs are reported for trials which give their results in the form of events per patient years—or at least insisting that commentaries make clear which method they are using.

An fuller explanation of the different methods can be viewed on www.suffolk-maag.ac.uk/ebm/pt-yrs&NNTs.html, with examples available for the UKPDS 38 trial (www.suffolk-maag.ac.uk/stats/cpukpds.html) and for the HOT trial (www.suffolk-maag.ac.uk/stats/cphot.html).

Kevork Hopayian, general practitioner.
The Surgery, Leiston, Suffolk IP16 4ES k.hopayian@tesco.net

John McGough, general practitioner.
The Surgery, Aldeburgh, Suffolk

Length of follow up is poorly reported

EDITOR—Smeeth et al raise interesting issues concerning the validity of reporting numbers needed to treat (NNTs) in systematic reviews that combine trials with varying periods of follow up.1 In such situations, only when the absolute treatment effect is constant over time can the NNT be correctly estimated from the reciprocal of the pooled absolute risk difference. By contrast, if a treatment has a constant relative effect over time, then within a single trial the NNT will decrease with increasing follow up.2 Similarly, we expect that the NNT will also vary among several similar trials with different lengths of follow up.

The reporting of length of follow up is often inadequate to assess whether the constant absolute risk model or constant relative risk model is the more appropriate in a given systematic review, or to make adjustments for length of follow up in the analysis. We assessed the quality of reporting of length of follow up in the systematic reviews published in the
Cochrane Library (Issue 1, 1998) that synthesised mortality outcomes. We excluded reviews in pregnancy and childbirth, where duration of follow up is typically not an issue. The 44 relevant systematic reviews that we found combined 306 trials. For 43% of the trials there was no mention of the duration of follow up in the published review.

To assess whether the cause was inadequate trial reporting or poor data abstraction we considered in more detail the 17 systematic reviews for interventions related to stroke, and compared the reporting of follow up in the reviews with that in the 103 trials on which they were based. We noted whether the reviewers had categorised the length of follow up as fixed (all participants studied for the same length of time) or variable (follow up summarised by mean, median, or range) or whether follow up was not stated. We found 93% agreement between the reviewers' abstractions and our own assessments, which suggests that poor reporting of trials is responsible for many of the omissions. These results support the results of other reviews of reporting of follow up in clinical trials and cohort studies. 3 4

Our findings suggest that, as so many trial reports omit mentioning length of follow up, in practice it may not be possible to adjust for length of follow up in a systematic review.

Roberto D'Amico, medical statistician. r.d'amico@icrf.icnet.uk

Jonathan J Deeks, medical statistician.

Douglas G Altman, professor of statistics in medicine. ICRF/NHS Centre for Statistics in Medicine, Institute of Health Sciences, Oxford OX3 7LF

We are grateful for the assistance of Hazel Fraser and the Cochrane Stroke Review group for allowing us access to copies of the 103 trials.

2. Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. BMJ (in press).

We are grateful for the assistance of Hazel Fraser and the Cochrane Stroke Review group for allowing us access to copies of the 103 trials.

NNT is a tool, to be used appropriately

EDITOR—There is much to agree with in the article on numbers needed to treat (NNT) by Smeeth et al. 1 But to use the word misleading in the title is in itself misleading. NNTs are a
huge advance on what we had before. Smeeth et al point out, as has been done previously, that for NNTs to be comparable they must define patients' condition and severity, the intervention, outcome, and duration, and perhaps other relevant issues. In saying that NNTs should reflect underlying baseline risk for an individual patient (or group of patients) they are restating a method described by Sackett et al.

The problem with their argument is that it is derived from examples of interventions used to prevent small effects in large numbers of patients. Most of us live in a medical world where we need interventions that produce large effects in small populations. In these circumstances, NNTs from meta-analysis are usually informative and seldom misleading.

Take acute pain as an example. Many high quality randomised, double blind, and placebo controlled clinical trials have been done over 50 years. For trials to be clinically valid patients have to have moderate or severe pain on entry. Pain is measured with standard scales over periods of 4-6 hours. Using the outcome of relief of at least half the pain over this time we have been able to calculate NNTs compared with placebo for a range of analgesic interventions (references on the BMJs website). NNTs are unaffected by pain model (dental or postoperative), pain measurement, duration (four or six hours), or reporting quality (given that trials are randomised and double blind).

Moreover, we have been able to use large amounts of data from individual patients and clinical trials to investigate the effect of chance on baseline and experimental event rates. Because individual clinical trials are set up to investigate the direction of treatment effect (treatment better than control), we need to know how much information is needed to overcome random effects in estimating the magnitude of the clinical effect of an intervention—or when an NNT becomes clinically valid.

NNT is a tool. Like any tool, when used appropriately it will be helpful and effective. What we have to do is to ensure that in any given situation we know what the rules are for using the tools correctly. Making swingeing oversimplifications from the same selected trials doesn't move us any further forward.

Andrew Moore, consultant biochemist. andrew.moore@pru.ox.ac.uk

Henry McQuay, clinical reader in pain relief. Pain Research, Churchill, Oxford OX3 7LJ


**Other related articles in BMJ:**

**EDUCATION AND DEBATE**

**Numbers needed to treat derived from meta-analyses—sometimes informative, usually misleading.**

Liam Smeeth, Andy Haines, and Shah Ebrahim

BMJ 1999 318: 1548-1551. [Full text]