Research methods

Evidence-based evaluation of benefits in therapeutic interventions: methodologically controlled and non-randomly assigned reflections on the number needed to treat

Gian Franco Gensini, Andrea A. Conti

Department of Critical Care Medicine and Surgery, University of Florence, Florence, Don Carlo Gnocchi Foundation, Florence, Italian Center for Evidence-Based Medicine, Florence, Italy

Key words:
Coronary artery disease; Evidence-based medicine; Therapy; Trials.

Background. **The number needed to treat (NNT), calculated as the reciprocal of the absolute risk reduction, is a parameter that provides quantitative information on the efficacy of therapeutic interventions. The introduction of this parameter was provoked by the demonstration that physicians were not too familiar with the percentage expressions of risk reduction, but preferred, in assessing the efficacy of a treatment in the context of a clinical trial, numbers that directly indicated patients.**

Description. **Although the results of the evolution in the concept of the NNT have been available for various years, the diffusion of such parameters is still poor. The frequency of explicit reporting of the NNT in the published randomized controlled trials of five major biomedical journals has been recently assessed. This review has shown that in 359 eligible papers, the NNT was recorded in only eight articles. This represents a major drawback of biomedical literature, since the NNT allows health operators to understand how much effort is needed to prevent a given event.**

Conclusions. **The NNT represents a measure of immediate and major clinical impact, and should always be reported, when appropriate, as an expression of the results of clinical trials. As a single number, it cannot stand alone as the sole justification for including a new treatment in the therapeutic armamentarium; still, it should be seen as an indicator by means of which the same treatment may be critically considered. It seems therefore appropriate to require that authors publishing results of clinical trials should provide readers with the information necessary to allow them to make responsible decisions; in this case, adopting the useful tool constituted by the NNT.**

(Ital Heart J 2003; 4 (2): 80-83)

Received September 23, 2002; revision received November 6, 2002; accepted November 11, 2002.

Address:

Dr. Andrea A. Conti

Dipartimento di Area Critica Medico-Chirurgica Università degli Studi Viale Morgagni, 85 50134 Firenze
E-mail: aabconti@yahoo.com

Background

The number needed to treat (NNT) and the number needed to harm (NNH)[1,2] are well-known parameters that provide quantitative information based on the type of intervention, its duration and the adverse outcome that the intervention itself attempts to prevent. The NNT is calculated as the reciprocal of the absolute risk reduction, the NNH as that of the absolute risk increase, and both must be accompanied by 95% confidence intervals (CI), i.e. the range of values, calculated on the basis of the sample data, that includes the real value of the studied parameter in 95 out of 100 cases[3].

In this paper we will focus our attention on the NNT, which was first described in 1988 by Laupacis et al.[4] in order to provide health operators with a quantitative expression of the number of patients who must be treated to prevent a given adverse outcome. The NNT represents the expected number of patients who must undergo an experimental treatment so as to prevent one additional adverse outcome event, or to realistically expect one additional beneficial outcome, compared with the expected event rates on control therapy[5].

Examples derived from a well-known randomized controlled trial could be useful to elucidate the calculation of the NNT. The CAPRICORN Investigators[6] have shown that adding carvedilol to the short-term management of myocardial infarction reduces, among others, the "all-cause mortality" and the "all-cause mortality or non-fatal myocardial infarction" in patients with left ventricular dysfunction. As a matter of fact, in the group allocated to carvedilol (975 patients), the percentage of patients reaching the endpoint "all-cause mortality", i.e. the absolute risk of this endpoint,

was approximately 12%, compared to the approximately 15% of patients (n = 984) assigned to placebo. The absolute risk reduction was therefore about 3%, and since the NNT is the reciprocal of the absolute risk reduction (1/3% = 100/3), the number of patients to be treated to prevent one additional all-cause death was around 33 (CI 16-244). With regard to the endpoint of "all-cause mortality or non-fatal myocardial infarction", since the absolute risk of the carvedilol patients was about 14% and that of the placebo ones was approximately 19%, the absolute risk reduction was 5%, and consequently the NNT was around 20 (in fact: 1/5% = 100/5 = 20), with CI comprised between 12 and 52.

The introduction of the NNT was prompted by the demonstration that physicians were not too familiar with the percentage expressions of risk reduction, such as the relative risk reduction and the absolute risk reduction, but preferred, in order to assess the efficacy of a treatment in the context of a clinical trial, numbers that directly indicated patients (to be treated with the studied therapy)[7]. Moreover, evidence exists that reporting the relative risk reduction alone may lead a reader to overestimate the effects of a given treatment[8].

Since the current best evidence of the efficacy of medical interventions derives from methodologically sound randomized controlled trials, a correct and adequate modality of reporting the quantitative results of such trials is mandatory for the proper evaluation of the therapeutic (but also the preventive) information deriving from them[9].

The number needed to treat and relative refined parameters

As far back as in 1999 Mancini and Schulzer[10] published in *Circulation* a paper in which they presented a further evolution of the concept of the NNT. In that paper the authors, recalling their previous proposal[11] of adding, to the consolidated NNT, the index "us" = "unqualified success", proceeded to calculate the clinical impact of this index by analyzing 11 clinical trials from the cardiovascular literature.

The authors define the NNTus as a measure of the degree of therapeutic effort necessary to prevent death in the patient without incurring in any serious, treatment-induced side effects; in other words, the ideal success all clinicians aspire to. It is our belief that a retrospective, methodologically controlled and non-randomly assigned comment on the potential clinical applications of the above parameter may be useful in highlighting the potentialities and limits of an evidence-based medicine quantitative approach.

The NNTus can be considered as a statistic refinement of the NNT concept, since, with its "evident" lexical absoluteness, it appears to aim at representing a more exact way of indicating the quantitative benefit parameter of the therapeutic approach under consideration. In fact, this treatment-specific index emerges from the comparison of the arms of trials, the therapeutic outcome under study, the duration of treatment necessary to reach the studied outcome, and the treatment-related adverse effects. It also furnishes really relevant quantitative indications to allow the making of decisions in the allocation of medical services and economic resources.

An example from the study of Mancini and Schulzer appears useful to better clarify the meaning of the NNTus. The Canadian Atrial Fibrillation Anticoagulation (CAFA) study[12] was one of the first clinical trials to evaluate the use of warfarin for the prevention of acute ischemic stroke in subjects with atrial fibrillation. The study patients (187 randomized to warfarin, 191 to placebo) were followed up for a mean of 2.5 years. The target international normalized ratio (INR) was between 2 and 3. In this study the desired outcome was represented by the prevention of a cluster of events including non-lacunar stroke, intestinal, renal or lower limb thromboembolic events, and fatal and intracranial hemorrhages. To prevent the embolic cluster (stroke and thromboembolic events) in 1 patient, approximately 59 patients had to be treated (standard error-SE about 72). The NNT was therefore 59. To achieve this objective without inducing also a fatal/major or minor hemorrhagic event, 65 patients had to be treated (the NNTus was therefore 65, with a SE of approximately 79). The NNTus derives in fact from the product of the reciprocal of the product of the absolute risk reduction multiplied by the probability of not having an adverse event due to the treatment. The absolute risk reduction is provided by the difference between the proportion of patients, in the non-treated population, who present with the embolic cluster (5.2%) and the corresponding proportion in the treated population (3.5%). The probability of not having an adverse event due to the treatment is given by the complement to 1 of the difference between the rate of adverse events in the treated group (18.5%) and the rate of adverse events in the control group (9.5%). In summary: 1/[(5.2-3.5%) × (1-9%)] = 1/[1.7 × 91%] = 1/(154.7/10000) = about 64.7. Therefore, the NNTus is usually higher than the NNT, since it "includes" the effort which is mandatory to prevent unfavorable outcomes in the absence of treatment-induced adverse effects.

Under-reporting of the number needed to treat in current medical literature

Although, as we have just seen, evolutions of the concept of the NNT have been available for various years, the diffusion of such parameters is still poor, especially in an era in which the quality of reporting for randomized controlled trials should improve in accordance with the Consolidated Standards of Reporting

Trials statement. Nuovo et al.[13] have recently evaluated the frequency of explicit reporting of the NNT (as well as of the absolute risk reduction) in the published randomized controlled trials of five major biomedical journals (*Annals of Internal Medicine, British Medical Journal, Journal of the American Medical Association, Lancet, New England Journal of Medicine*). For each journal they have evaluated 4 non-consecutive years, and all issues of each journal for each year have been reviewed manually. Among 359 eligible papers, the NNT was recorded in just 8 articles, 6 of which dated 1998 (the most recent among the 4 years examined). British and US journals did not show significant differences in the frequency of reporting results with the NNT, and even the *British Medical Journal* – one of the biomedical journals paying more attention to evidence-based medicine – did not seem better than the others. This appears to be quite disappointing since the NNT has a number of merits and potentialities to be remembered.

Merits and limits of the number needed to treat

As a matter of fact, the NNT allows health operators to understand how much effort is needed to prevent one (unwanted) event, thus permitting comparisons with the quantity of effort needed to prevent the same or other events in subjects with other disorders or affections[14]. Moreover, the NNT indicates the efficacy by incorporating the baseline risk without therapy and the risk reduction of therapy, and it is therefore currently recommended especially for binary and survival time data[15].

Just as for every expression of the efficacy of a treatment, and in the light of a real evidence-based approach, the NNT too has limits. These should be fully acknowledged by health professionals to put into perspective its real and relevant value. With special reference to the NNTus, such limits include its restricted usefulness, in the comparison of different diseases, to those sole cases in which the length of therapy and the outcomes are similar, in view of the fact that its estimation is punctual rather than dynamic. The need for further methodological research on its application and validation in clinical settings other than the cardiovascular one, as well as in sets of trials with different methodological designs, should also be signaled[16]. With regard to the NNT, it should be said that its use is less appropriate when the interventions being compared have effects over different periods, or are applied to different populations, or when the results of a study with patients at one baseline risk are compared to those of patients with a different risk[17]. Ebrahim and Smith[18] have demonstrated that the reliability of the NNT is much higher when it is derived directly from a clinical trial than when its value has been calculated on the basis of pooled absolute risk differences in meta-analyses; in the latter situation it should be considered with particular caution.

Conclusions and recommendations

In the light of a broad evidence-based, and not only NNT-based, approach, and in spite of its limitations, the NNTus is emerging as interesting in its potential relevance, and is proposed as providing a reliable color picture of the enormous clinical effort necessary to achieve an objective balance of benefit and harm, a balance that clinical practice daily tries to pursue with the great aid of evidence-based medicine and associates[19].

The NNT represents, on its part, a measure of immediate and major clinical impact, and should be reported – when appropriate – as an expression of the results of clinical trials. Since authors do not always provide complete information for this calculation, it cannot even be left to be worked out by the readers, and this is a current major limitation of randomized controlled trials[20]. As a sole number, it certainly cannot stand alone as the single justification for including a new treatment in the therapeutic armamentarium, but it certainly should be seen as an indicator with which to critically consider the same treatment.

Today it therefore seems appropriate to ask that authors publishing the results of clinical trials provide the readers with the adequate information to allow them to make responsible decisions, also adopting the useful tool constituted by the NNT, a clinically useful "yardstick" of the effort required to achieve an additional beneficial outcome or prevent an additional unfavorable one with a specific health treatment[21].

References

1. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine. How to practice and teach evidence-based medicine. New York, Edinburgh, London: Churchill Livingstone, 1997.
2. Greenhalgh T. How to read a paper. The basics of evidence based medicine. London: BMJ Publishing Group, 1997.
3. Everitt BS. The Cambridge Dictionary of Statistics in the Medical Sciences. London: Cambridge University Press, 1995.
4. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med 1988; 318: 1728-33.
5. Walter SD. Number needed to treat (NNT): estimation of a measure of clinical benefit. Stat Med 2001; 20: 3947-62.
6. The CAPRICORN Investigators. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: the CAPRICORN randomised trial. Lancet 2001; 357: 1385-90.
7. Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? Ann Intern Med 1992; 117: 916-21.
8. Elting LS, Martin CG, Cantor SB, Rubenstein EB. Influ-

ence of data display formats on physician investigator's decisions to stop clinical trials. BMJ 1999; 318: 1527-31.
9. Chatellier G, Zapletal E, Lemaitre D, et al. The number needed to treat: a clinically useful nomogram in its proper context. BMJ 1996; 312: 426-9.
10. Mancini GBJ, Schulzer M. Reporting risks and benefits of therapy by use of the concepts of unqualified success and unmitigated failure. Applications to highly cited trials in cardiovascular medicine. Circulation 1999; 99: 377-83.
11. Schulzer M, Mancini GBJ. "Unqualified success" and "unmitigated failure": number-needed-to-treat-related concepts for assessing treatment efficacy in the presence of treatment-induced adverse events. Int J Epidemiol 1996; 25: 704-12.
12. Connelly SJ, Laupacis A, Gent M, Roberts SR, Cairns JA, Joyner C. Canadian Atrial Fibrillation Anticoagulation (CAFA) Study. J Am Coll Cardiol 1991; 18: 349-55.
13. Nuovo J, Melnikov J, Chang D. Reporting number needed to treat and absolute risk reduction in randomised controlled trials. JAMA 2002; 287: 2813-4.
14. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. BMJ 1995; 310: 452-4.
15. Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomised trials: explanation and elaboration. Ann Intern Med 2001; 134: 663-94.
16. Wu LA, Kottke TE. Number needed to treat: caveat emptor. J Clin Epidemiol 2001; 54: 111-6.
17. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guide to the medical literature, IX: a method for grading health care recommendations. JAMA 1995; 274: 1800-4.
18. Ebrahim S, Smith GD. The "number needed to treat": does it help clinical decision making? J Hum Hypertens 1999; 13: 721-4.
19. Isaacs D, Fitzgerald D. Seven alternatives to evidence based medicine. BMJ 1999; 319: 1618.
20. Nuovo J. Interpreting the number needed to treat. (letter) JAMA 2002; 288: 7.
21. Cordell WH. Number needed to treat (NNT). Ann Emerg Med 1999; 33: 433-6.