A medication or method of treatment is more successful the more often the desired objective is achieved by intervention. This is expressed as the "number needed to treat" (NNT). Ideally, this number should be one. Only then are patients who benefit from a treatment also the only ones who are exposed to the risk of undesirable effects. In acute diseases, methods of treatment are unlikely to endure for long unless they produce successful results in at least every third or fourth patient (corresponding to an NNT of 3 or 4). In the preventive treatment of asymptomatic patients, as well as in the management of seriously ill persons, it cannot be stated in each specific case whether the treatment has, e.g., been able to prevent a stroke or prolong life.

Some of the pitfalls involved in assessing NNT and relative risk are described and, taking studies on antihypertensives as an example, the relationship between background risk and NNT is explained. A proposal for an extended formulation of NNT is also presented along with ways of reducing NNT by more thorough diagnostic evaluation.

**NNT versus relative risk reduction**

In randomized controlled trials and metaanalyses, the risks (or failure rates) of treatments are estimated and compared. If, with the same number of patients in two relatively large compared groups, 100 of the untreated patients but only 80 of the treated patients become ill, this represents a relative risk of 0.8. This can also be expressed as a 20% risk reduction, which sounds more impressive. This figure of 0.8, however, requires further clarification: a risk reduction of 20% causes many people to assume that in the absence of treatment everyone falls ill, and that if treatment is given, approximately every 5th person is spared this fate (1, 2). This fallacy is commonly used in pharmaceutical advertising.

How can we assess the implications of absolute risk with and without treatment? In the example quoted with a relative risk reduction (RRR) of 20%, there might, e.g., be 50% (100 of 200) in untreated subjects and 40% (80 of 200) in treated subjects, or there might be 5% (100 of 2 000) and 4% (80 of 2 000). In the latter case it would be necessary to treat not only 5 patients to prevent an event, but 10 or even 100. NNT appears to provide the best information about treatment efficacy. It is widely used and accepted as a means of comparing pharmaceutical and non-pharmaceutical treatments (3), and in comparing different sub-groups in case control studies. It is often assumed that a smaller NNT implies a more effective treatment. This is only the case, however, if the comparisons are admissible, as explained below.

---

**SUMMARY**

**Introduction:** Aggressive pharmaceutical industry advertising, as well as authors and journals, often emphasize a relative risk reduction. These results can be adjusted by the number needed to treat (NNT). **Methods:** Simple mathematical considerations and selective literature review. **Results:** Without any further information, NNT can be misleading. In controlled clinical trials, the NNT identifies the superior treatment correctly if the background risks are comparable. From this background risk the smallest possible NNT can be calculated. In addition, NNT must be adjusted for duration of treatment. The difficulties in assessing outcomes are exemplified by studies on antihypertensives. **Discussion:** The ratio between the smallest possible NNT and the calculated NNT should be mentioned (factor of dilution). Not only therapeutic efficacy but also diagnostic accuracy can reduce the NNT.

**Key words:** drug efficacy, clinical evaluation, study, health economics, number needed to treat.
Relationship between efficacy and NNT

An ideal medication could invariably prevent a certain disease event which occurs, e.g., in 2% of a population. The NNT for treatment with this medication is 50. Since it is not known which two persons are threatened by the disease, everyone has to be treated. With a 5% risk of developing disease, the NNT would decrease to 20 with the same treatment (diagram 1a). Only if the disease event occurred in all 100 patients in the absence of treatment (e.g. death due to lack of dialysis in patients with end-stage renal failure) is the NNT 1.

When only the NNT is considered, equivalence of treatments is easy to fake: a patient population with a disease risk is treated with a medication which prevents the disease event only in every second patient. By the end of the treatment, 5% in the active group and 10% on placebo would become ill, yielding an NNT of 20 (diagram 1b). This is the same NNT as for a much better medication (diagram 1a), when used in patients with a 5% risk.

The NNT increases with a decrease both in the efficacy of treatment and of the background risk. The following comparison of the results of two placebo controlled studies shows how different background risks can lead to different NNTs. Two studies on antihypertensives in elderly subjects yielded NNTs of 175 and 71 for the prevention of stroke. In the SYST-EUR study investigators tested a calcium channel blocker, and in the STOP study a diuretic and a beta blocker (10). In the first study the stroke risk was 1.39 per 100 patient years and decreased to 0.82, in STOP 3.1 per 100 patient years and decreased to 1.7. The smallest possible NNT was 72 in SYST-EUR, but in the second study 32, and therefore the chances of a small NNT were much greater in the second than in the first study. The ratio of actual to minimum possible NNT was approximately equal with 2.4 and 2.2. Expressed differently: for the prevention of stroke, a diuretic or beta blocker therapy cannot be regarded as superior to treatment with the calcium channel blocker despite the low NNT in the STOP study.

NNT in studies on antihypertensives

This does not mean that NNT is valueless. The examples chosen merely demonstrate that NNT is dependent on the risk to which patients are exposed if untreated, i.e. the background risk. Its practical significance for NNT is seen, e.g., in other studies with antihypertensives (diagram 2). One study evaluated trials published up to 2003 (8) which included composite analyses of severe cardiovascular events. As an aid to comparing the calculated NNT, they have been standardized here to one year with a largely constant increase in therapeutic benefit over time. NNTs show a wide range of variability; a semi-logarithmic scale was therefore chosen. The smallest possible NNT (NNT_{min}), derived by calculation (box 1) from the risk in the placebo group, are also shown for each study.

A treatment in which the actual and the smallest possible NNT are identical is intuitively regarded as particularly effective. Diagram 2 shows that this is sometimes almost achieved, although in the medium risk range, the difference between actual NNT and the smallest possible NNT is relatively constant. This difference corresponds, because of the logarithmic presentation, to a factor by which the smallest possible NNT can be multiplied to obtain the actual NNT.

BOX 1

Calculation of the dilution factor of an NNT from study results

\[
\begin{align*}
RRR &= 100 \times (1 - \text{event rate}_{\text{with therapy}}/\text{event rate}_{\text{without therapy}}) \\
ARR &= \text{event rate}_{\text{without therapy}} - \text{event rate}_{\text{with therapy}} \\
NNT &= 100/ARR \\
NNT_{\text{min}} &= 100/(\text{event rate}_{\text{without therapy}}) \\
NNT/NNT_{\text{min}} &= \text{dilution factor} \\
100/RRR &= \text{dilution factor} \\
NNT/\text{dilution factor} &= NNT_{\text{min}}
\end{align*}
\]

Event rates in %
a), b) and c) stand for prospective controlled cardiovascular studies. Black squares symbolize cardiovascular events. When considering exclusively NNT in a) and b), an apparent equivalence of medication 1 and 2 is found, although medication 1 prevents all complications, and medication 2 only prevents every second complication. In c) drug treatment is not initiated or is discontinued, e.g. in favor of surgical treatment, because diagnostic evaluation did not disclose an elevated risk when other factors were taken into account (unshaded square) or because the additional risk was eliminated by alternative treatment (crosshatch) or can be sufficiently reduced by general, nonpharmacological means (unshaded circle). As a result, the group of subjects permanently treated with antihypertensives is reduced from 100 to 80, and the NNT also decreases due to administration of medication 2.
NNTs are now compared for various studies which do not take into account background risk and differences between the patient groups and the study conditions. Treatment A and B were given in patients with an approximately equal risk. The higher NNT of treatment A indicates its inferiority compared to treatment B. In studies C and D, the patient risk differed by a factor of 5. Although the NNT was approximately equal in both studies, treatment D is very likely to be poorer. The decision is more difficult when the NNT is higher for B compared to E with a very similar difference in relation to the respective smallest possible NNT. Because E was used in patients with an approximately 1.5 fold risk compared with B, B and E should be compared in a new controlled study, i.e. in patients with equivalent risk. Likewise, it is probably unjustified to reject the treatment used in study C whose NNT was above an arbitrary level, in this case 100 per year. The NNT for C would probably markedly decrease with a higher baseline risk. For this reason, the efficacy of a treatment should not be judged solely on the basis of NNT, but with reference to the ratio of actual to smallest possible NNT. This ratio is a mathematical equivalent of the relative risk reduction.

Extended formulation of NNT and limits of evidence value

The examples demonstrate that stating the NNT alone is not sufficient to compare the efficacy of different treatments, even if the NNT, as in diagram 2, have been standardized to a certain time period. However, this is not true in every case. The ratio of smallest possible NNT to actual NNT can be more informative. This figure, which may be referred to as dilution factor, should be simultaneously reported. For diagram 1b this would yield the following:

\[
\text{NNT: } \frac{100}{10–5} = 20, \quad \text{smallest possible NNT (NNTmin): } \frac{100}{10} = 10, \quad \text{dilution factor: } \frac{20}{10} = 2.
\]

The extended formulation of the NNT therefore becomes \(20/2\) (box 1); this shows, among other things, that \(20/2\) can also be read as a division, the resolution of which yields the smallest possible NNT. The more closely the figure after the slash approaches 1, the better the efficacy of a therapy, because the actual NNT determined in a study approaches the smallest possible NNT. The efficacy of a procedure with an NNT of 10 000/1 (e.g., a vaccination campaign) is therefore beyond doubt. Its use depends mainly on the benefit-harm ratio and the economic context.

Some considerations have been omitted, including (4, 5)
- confidence intervals for the NNT and the dilution factor
- the need for a statistically significant study as a precondition for NNT calculation
- standardization of the NNT to a specific time interval where the course is nonlinear.
The confidence interval for NNT is calculated from the confidence interval of the absolute risk reduction (ARR). A statistically insignificant study result includes an absolute risk reduction (ARR) of zero. This would imply an infinite (6), or even negative NNT. We also omit detailed discussion of the problem that the efficacy of treatments is often not comparable in different patient populations and under different study conditions. The term NTN ("number needed to treat needlessly") is also used for "NNT-1". The information content of NTN is low. It is easily overlooked that the absent need for treatment relates only to one specific hard endpoint in a certain time period. In studies with positive outcomes, pathophysiological processes are beneficially influenced by the treatment during this period in some of the patients who are seemingly treated unnecessarily, and the avoidance of hard endpoints may lie beyond the study period. The NTN may appear exaggerated, while NNT is rather a cautious or conservative benefit descriptor.

Smaller NNT with clearer diagnostic work up

The economic significance of NNT as a basis for cost-effectiveness analyses cannot be considered here. In preventive strategies the question arises as to whether constantly new and often expensive medications result in the smallest possible NNT. It may be that we are like the proverbial donkey pursuing the carrot suspended at a fixed distance in front of its nose. Often a strategy aimed at reducing the smallest possible NNT might be more effective.

Diagram 1c demonstrates that the NNT can be reduced at the same medication strength. To do this, it is necessary to identify which patients benefit little, which not at all, and which markedly from a treatment. This improves diagnostic accuracy and reduces NNT. Approaches adopted in oncology and infectious diseases may be regarded as groundbreaking in this respect (7). Diagnostic clarification need not necessarily precede treatment, but can take place in parallel.

Almost all recommendations for the treatment of hypertension and hyperlipidemia contain advice on cardiovascular risk stratification. It is proposed that patients should be excluded from long-term medication who have exceeded the normal limits for a single risk factor, possibly transiently, against a background of normal cardiovascular risk. These recommendations have been implemented incompletely and have proved inefficient as a means of reducing NNT. The stratification guidelines lack consistency. Risk calculations differ according to the epidemiological data survey from which a risk score is derived. Furthermore, some researchers tend, depending on the main focus of their research, to include blood pressure or lipid values quantitatively but the other parameters only qualitatively in the risk assessment. This compromises the accuracy of the risk estimate.

Diagnostic procedures, especially those involving more recent noninvasive methods, generate costs. Noninvasive investigation, however, is usually more helpful than futile drug treatments given over a period of years. However, the treating physician is unable to off-set the cost of diagnosis costs against costs saved via treatment. There is also the risk that physicians who fail to prescribe or discontinue prescribing are perceived by some patients...
as incompetent and, in the current climate of limited prescribing budgets in Germany, motivated by profit (box 2).

The evidence base suggests the conclusion that NNT can be reduced by more thorough diagnostic evaluation and therapy modifications. As long as there remains an evidence gap in this area, plain common sense, as so often in medicine, will have to suffice.

Acknowledgment
The author thanks Prof. Dr. rer. nat. Gerhard Osius, Institute of Statistics of Bremen University, for providing suggestions and reviewing the manuscript.

Conflict of Interest Statement
The author declares that no conflict of interests exists according to the Guidelines of the International Committee of Medical Journal Editors.

Manuscript received on 6 March 2006, final version accepted on 25 August 2006.

Translated from the original German by mt-g.

REFERENCES

Corresponding author
Prof. Dr. med. Manfred Anlauf
Friedrich-Plettke-Weg 12
27570 Bremerhaven, Germany
manfred.anlauf@t-online.de