An Inadvertent but Explicable Error in Calculating Number Needed to Treat for Reporting Survival Data

Sudhir K. Bowry,* Volker Schoder,† and Christian Apel‡

*EMEALA Medical Board, †Department of Clinical Research, and ‡Department of International Strategic Development, Fresenius Medical Care Deutschland GmbH, Bad Homburg, Germany


Drawing one’s own conclusions from the findings of a clinical trial is based on knowledge of facts as well as the quality and source of the evidence presented. Statistical methodology that fits the study design chosen to answer research questions defends the relevance and credibility of the acquired data. Even then, grasping the value of research is all too often daunting, unless one is well versed with elements of statistics applied to clinical trials. When an unfortunate statistical slip creeps into research papers, judging the validity of trial results becomes more difficult. Prompt correction and explanation of the error is the responsible course of action, but more importantly, it should serve to prevent the error from being repeated in future clinical research and its interpretation.1,2

Such an innocent error was noticed in a recent randomized controlled trial (RCT) evaluating the use of high-volume hemodiafiltration therapy in prolonging survival of dialysis patients with end stage CKD.3 In assessing the clinical benefits of this treatment modality, Maduell et al.3 applied “number need to treat” (NNT) or number of patients who need to be treated to avoid one additional event, which is used to express comparisons of success or failure of different therapies.3

Proposed originally by Laupacis et al.4 in 1988 as a more meaningful clinical measure of the consequences or value of a treatment, the main appeal of the NNT is that it is easier than other statistical parameters to understand and interpret the value of clinical trials. It allows a straightforward comparison of therapy procedures to convey both statistical and clinical significance to the physician. NNT has, thus, become an intuitive and widely used measure of treatment benefit derived from the results of RCTs with a binary outcome (e.g., success/failure, yes/no, or complications present/not present). Relative risk, absolute risk reduction (ARR), odds ratio, and hazard ratio (HR) are among the other more common statistical measures of comparing intervention and control groups in a trial. The basic mathematical language required to describe each of these familiar entities as well as the limitations and interchangeability between them have been summarized by Scott.5

In its original form, NNT is an epidemiologic measure to present data where two treatments are compared with respect to incidence rates of an unfavorable event.6 NNT is simply calculated as the reciprocal of the ARR, which is the difference between the absolute risk of an event in the intervention group (treatment A) and the absolute risk in the control group (B). For example, assuming that five events occur among 100 patients treated in group A and seven events occur among 100 patients treated in group B, then the ARR is 0.07–0.05=0.02. Thus, NNT is 1/0.02=50.

However, many clinical trials in nephrology are designed to assess the benefit of an intervention in terms of survival (i.e., the primary outcome is time to an event [death]). In this case, NNT will vary according to the length of the follow-up in the study; hence, there is no single NNT, but rather, it can be calculated at any time point in the study follow-up. In their publication in 1999, two renowned medical statisticians, Altman and Andersen,7 addressed this issue from a statistical perspective to show how to calculate NNT depending on the type of survival data available.

Where only a simple survival analysis has been performed (Kaplan–Meier survival curves available), ARR is expressed as the difference in survival rates at time t [S_A(t) − S_B(t)], and NNT is, again, its reciprocal:

\[ NNT(t) = \frac{1}{S_A(t) - S_B(t)}. \]  [1]

If, however, an estimate of the HR is additionally available (as was the case for the trial by Maduell et al.), Altman and Andersen give an alternative way to calculate NNT:

\[ NNT(t) = \frac{1}{(S_B(t))^{HR} - S_B(t)}. \]  [2]

where \( S_B(t) \) denotes the Kaplan–Meier survival probability in the control group at time t and \( HR \) refers to the HR (test group versus control) that can be directly taken from a Cox regression analysis.
Equations 1 and 2 are required to point out and explain the inadvertent but perhaps not entirely inexplicable error that arose in the paper by Altman and Andersen.7 That was then carried forward by Maduell et al.3 Considering that the terms survival \( S \) and mortality \( M \) are often interchangeably used in scientific literature and that \( S(t) = 1 - M(t) \), replacing survival probabilities \( S(t) \) by probabilities of mortality \( M(t) \) in Equation 1 would not alter the NNT value. However, in Equation 2, it is mandatory to use only survival probabilities, because replacing \( S(t) \) by \( M(t) \) would result in altogether different values for NNT. Simply put, in the landmark paper by Altman and Andersen,7 the example calculation from an RCT shown in the paper, mortality rates appear instead of survival rates in Equation 2. That is, Altman and Andersen applied in their example using Equation 2 a mortality rate of 0.33, whereas they should have used a corresponding survival rate of 0.67.

Maduell et al.3 derived NNT for 1-year survival of 9.8 using the example calculation of Altman and Andersen by applying a 1-year mortality rate of 10.3% in Equation 2. If 1-year survival rate of 89.7% would have been used instead, the correct NNT value is 33.2. Similarly, NNT values for 2- and 3-year survival change from 7.7 to 15.1 and from 7.7 to 10.5, respectively. It is important to document that this recalculation does not in any way undermine the key findings and indeed, the value of the Estudio de Supervivencia de Hemodiafiltración On-Line trial, because NNT was not the primary outcome measure of the study.

Bringing such methodological errors to the attention of the readership is in the interests of scientific candor and correct reporting and interpretation of RCTs.2 To our knowledge, the aforementioned discrepancy in the paper by Altman and Andersen7 has not been documented in any publication. NNT is a valuable tool that facilitates judicious use of available evidence for clinical decision-making and is even incorporated in the recommendations of the Consolidated Standards of Reporting Trials guidelines that strive to improve the quality of reporting of RCTs.8 The term has been gaining importance in all medical fields, including nephrology, for the publishing and evaluation of results of clinical trials.3,9–11 Scientific errors, experimental or statistical, may be inevitable and sometimes, go undetected. Recognition of the inadvertent glitch in the work by Altman and Andersen7 would enhance the applicability of NNT to critically appraise clinical evidence.

ACKNOWLEDGMENTS

The authors dedicate this article to Dr. Frank Thalau, who before he sadly passed away, provided his expertise during our discussions on number needed to treat.

DISCLOSURES

S.K.B., V.S., and C.A. are employees of Fresenius Medical Care Deutschland GmbH, Bad Homburg, Germany.

REFERENCES

7. Altman DG, Andersen PK: Calculating the number needed to treat for trials where the outcome is time to an event. BMJ 319: 1492–1495, 1999