by 25% to 33% overall. Of more importance is that the proportion of children with extended illness—that is, those who do not experience signs of improvement until after the fourth day of illness—was reduced from 20% to 3% in placebo controlled clinical trials.[1-3] In other words, those who need it most benefit most. What price tag can we attach to that?

D J Paynton's family probably did not experience the usual situation of secondary cases in the family being more severe than the index case. My family was less fortunate. One winter before oral acyclovir was available my three children had chickenpox one after the other in painfully slow succession. My son, the tertiary case, was nearly admitted to hospital because of central nervous system complications that were, in retrospect, due to a concomitant infection with influenza A virus.

The service implication of an additional treatment option is a valid concern. My suggestion would be to set up a triage system so that only questionable cases need be seen by a practitioner. In my experience, the parents or guardians are reliable at recognising chickenpox, epecially after they have suffered through the index case.

HANK BALFOUR
Professor

Box 437 UMHC,
15-144 PWB Harvard Street at East River Parkway,
Minneapolis,
MN 55455, USA

1 Balfour HH Jr, Kelly JM, Suarez CS, Heussner RC, Englund JA, Crane DC, *et al.* Acyclovir treatment of varicella in otherwise healthy children. *J Pediatr* 1990;**116**:633-9.
2 Dunkle LM, Arvin AM, Whitley RJ, Rotbart HA, Feder HM, Feldman S, *et al.* A controlled trial of acyclovir for chickenpox in normal children. *N Engl J Med* 1991;**325**:1539-44.
3 Balfour HH Jr, Rotbart HA, Feldman S, Dunkle LM, Feder HM Jr, Prober CG, *et al.* Acyclovir treatment of varicella in otherwise healthy adolescents. *J Pediatr* 1992;**120**:627-33.

Number needed to treat

Absolute risk reduction may be easier for patients to understand

EDITOR,—I strongly support the use of absolute risk rather than relative risk in clinical decision making[1] as the higher the patient's risk before treatment the greater the benefit of treatment. In New Zealand clinical guidelines for the management of mildy raised blood pressure,[2] the detection and management of dyslipidaemia,[3] and the use of hormone replacement therapy[4] have explicitly used the concept of absolute risk as a decision making tool. This concept, however, has been expressed as an absolute risk reduction rather than "numbers needed to treat." This absolute risk reduction is expressed in terms of the numbers of events prevented per 100 people treated.

If long term hormone replacement therapy is used as an example we can see how it is possible to weigh up the benefits and risks of treatment with this approach. For this example I have assumed from clinical trials that long term treatment with oestrogen alone reduces hip fractures by 25% and coronary heart disease by 35% and increases breast cancer by 25%.[5] Giving hormone replacement therapy for 15 years to 100 New Zealand women from the age of 50 would be expected to prevent roughly two myocardial infarctions, cause about one case of breast cancer, and confer little or no benefit on hip fracture. After 25 years (to age 75) the net difference in events would increase: for every 100 women given hormone replacement therapy there would be six fewer myocardial infarctions, one fewer hip fracture, and one to two additional cases of breast cancer. Benefits for hip fracture become greater after 75 years of age, when the risk of hip fracture increases steeply. This model may be modified for patients at different baseline risk.

As a general practitioner, I helped to develop the guidelines on hormone replacement therapy[4] and have informally evaluated how useful general practitioners find the concepts of numbers needed to treat and absolute risk reduction. Twenty general practitioners were asked in a pilot study of future guidelines on the management of mildly raised blood pressure which approach they would prefer in clinical practice. All found tables of absolute risk reduction, expressed in terms of the number of events prevented per 100 people treated, more useful than numbers needed to treat.

The concept of numbers needed to treat is useful for clinicians, policymakers, and economists illustrating the clinical and resource implications of treating patients with different baseline risks. When weighing up the benefits and risks of treatments with individual patients I have found that patients are better able to personalise information about treatment when the concept of absolute risk reduction is used. To promote the understanding and use of absolute risk in the clinical setting I recommend the use of tables of absolute risk reduction.

DIANA NORTH
Bristol-Myers Squibb research fellow

Department of General Practice,
Auckland School of Medicine,
Private Bag 92 019,
Auckland,
New Zealand

1 Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452-4. (18 February.)
2 Jackson R, Barham P, Bills J, Birch T, McLennan L, MacMahon S, *et al.* Management of raised blood pressure in New Zealand: a discussion document. *BMJ* 1993;**307**:107-10.
3 Mann JI, Crooke M, Fear H, Hay DR, Jackson RT, Neutze JM, *et al.* Guidelines for detection and management of dyslipidaemia. *NZ Med J* 1993;**106**:133-41.
4 MacMahon S, North D. *Hormone replacement therapy. A consensus development conference report to the National Advisory Committee on Core Health and Disability Support Services.* Wellington: Core Services Committee, 1993.
5 Grady D, Rubin SM, Petitti DB, Fox CS, Black D, Ettinger B, *et al.* Hormone replacement therapy to prevent disease and prolong life in postmenopausal women. *Ann Intern Med* 1992;**117**:1016-37.

Risk measures expressed as frequencies may have a more rational response

EDITOR,—Richard J Cook and David L Sackett propose that the number needed to treat (to avoid an adverse event) should be used as a measure of the efficacy of treatment in reducing medical risks.[1] This measure has advantages over probability measures (for example, reduction in relative risk) as it conveys both statistical and clinical significance.

A treatment that reduces deaths by 50% sounds better than one that reduces deaths by 5%, yet the latter treatment might be more valuable than the former: reducing a tiny risk by 50% might be trivial relative to reducing a large risk by 5%. The number needed to treat reflects the incidence and is more relevant for medical decision making.

Studies of medical decision making support use of this measure. Several studies show that experts —including medical clinicians—have great difficulty reasoning with probabilities.[2] For instance, Eddy asked 100 physicians questions of the following type. The prevalence of breast cancer is 1% (in a specified population). The probability that the result of mammography is positive if a woman has breast cancer is 79% and 9·6% if she does not. What is the probability that a woman with a positive result actually has breast cancer?[3]

Eddy reports that 95 physicians estimated the probability P (cancer and positive result) to be about 75%; the correct probability is only about 8%. Dawes reports a surgeon in the United States performing preventive mastectomies on the basis of this faulty logic.[4]

Nevertheless, the same problems presented by use of frequencies rather than probabilities are solved relatively easily. Gigerenzer reviewed several studies that show a dramatic improvement in reasoning with probabilities if they are converted into frequencies. For instance, we can change the example above as follows. Imagine 100 people (think of a 10×10 grid). We expect that one woman has cancer and a positive mammogram. Also we expect that there are 10 more women with positive mammograms but no cancer. Thus we expect 11 people with positive mammograms. How many women with positive mammograms will actually have cancer?

With frequencies you immediately "see" that only about one out of 11 women with a positive result will have cancer. Although staff of Harvard Medical School have difficulties with the probability version—most give wrong answers[5]—most undergraduates readily provide the correct answer to similar problems constructed with frequencies.[5]

Psychological research suggests that measures of risk communicated in terms of frequencies rather than probabilities will be more readily understood and rationally responded to.

PETER AYTON
Lecturer in psychology

Department of Psychology,
School of Social Sciences,
City University,
London EC1V 0HB

1 Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452-4. (18 February.)
2 Ayton P. On the competence and incompetence of experts. In: Wright G, Bolger F, eds. *Expertise and decision support.* London: Plenum, 1992:77-105.
3 Eddy DM. Probabilistic reasoning in clinical medicine: problems and opportunities. In: Kahneman D, Slovic P, Tversky A, eds. *Judgement under uncertainty: heuristics and biases.* Cambridge: Cambridge University Press, 1982:249-67.
4 Dawes RM. *Rational choice in an uncertain world.* Cambridge: Cambridge University Press, 1988.
5 Gigerenzer G. Why the distinction between single event probabilities and frequencies is important for psychology (and vice-versa). In: Wright G, Ayton P, eds. *Subjective probability.* Chichester: Wiley, 1994:129-61.

Computer software that can calculate confidence intervals is now available

EDITOR,—Richard J Cook and David L Sackett usefully show applications of the number needed to treat statistic for clinical decision making.[1] Confidence intervals are an important aid to meaningful inference from this statistic.[2] Mention in the paper of the construction of confidence intervals for the number needed to treat is limited to their direct inverse relation with confidence intervals for the absolute risk reduction. In statistical terms, the absolute risk reduction is analogous to the comparison of unpaired proportions. A widely published formula for approximate confidence intervals for the difference between unpaired proportions is often used to calculate confidence intervals for absolute risk reduction and thence number needed to treat.[3] With large numbers, such as those quoted by Cook and Sackett, this approximation is acceptable, but with smaller numbers it is unreliable.[4]

This problem can be overcome with computer software that uses robust iterative methods to construct these confidence intervals. Arcus Pro-Stat version 3.23 provides "near exact" confidence intervals for relative risk, relative risk reduction, absolute risk reduction, number needed to treat, and other statistics used in clinical epidemiology (Medical Computing, Aughton, Lancashire). Any reader who needs computer software for these calculations can obtain a royalty free copy of a limited version of Arcus Pro-Stat for IBM compatible computers by contacting me.

IAIN E BUCHAN
Honorary lecturer

Department of Medicine,
University of Liverpool,
PO Box 147,
Liverpool L69 3BX