

- 10 This interval offers:
 11 This interval does not offer:
 12 Where the interval ranges broadly and includes 1.0, it suggests:
 13 The following suggestion is designed to block rather than assist your understanding of the issues:

Options

- a an arbitrary dichotomisation of the data.
 b upper and lower limit values that are equally likely.
 c the range within which the true value lies.
 d a modern method to test for statistical significance.
 e the experiment was not sufficiently powered to achieve a result.
 f the experiment was not sufficiently discriminating to achieve a result.

EMI correct matchings

1	2	3
a, b, d, e, f	d, e, f	a, b, f
4	5	6
none	e	a, b, c
7	8	9
a, b, c, f	d, e	h
10	11	12
a, b	c, d	f
13		
e		

Providing the best available evidence

INVITED COMMENTARY ON... THE ANTIDEPRESSANT TALE

John Geddes

Abstract It is well recognised that the methods used to present the results of clinical research can substantially affect their interpretation. The critical appraisal of research articles requires the clinician to go beyond the 'spin'. The task is to summarise research studies in a way that assists the patient to make informed decisions guided by the best available evidence, their own preferences and the clinician's expertise.

Using the data on selective serotonin reuptake inhibitors (SSRIs), David Healy illustrates the difficulty of excluding the possibility of an association between a therapy and a serious, but uncommon, event such as suicide (Healy, 2006, this issue). The danger of relying on a statistical significance test to exclude clinically significant effects of treatment is, of course, well known but none the less worth rehearsing using the experience with the SSRIs. The basic issues are well explained with several examples by Altman & Bland (1995) and it is worth remembering the title of their article: 'Absence of evidence is not evidence of absence'. The commonly made error of interpreting a non-significant hypothesis test as meaning that there is no association is known as a type II error.

Although such a mistake may simply be due to lack of statistical expertise, it is also a way of 'spinning' the results in the direction that the authors would prefer them to go. Clinically, the events we are concerned with here will often be unexpected adverse events and the primary randomised evidence will usually have insufficient power to confirm or exclude an association reliably.

This is an increasingly important issue in clinical practice because service users are rightly demanding better information on the potential risks of treatments. So how should the clinician interpret data on risk? Evidence-based practice makes use of tactics derived from clinical epidemiology to identify the most robust research evidence and critically appraise it

for its validity and applicability (Sackett *et al*, 1996). David Healy mentions evidence-based practice but does not really do it justice because he does not refer to the standard evidence-based practice approach to dealing with evidence on the harmful effects of drugs (or other disease risk factors). It may be helpful to review briefly how to use quantitative estimates of harm, derived from an appropriate study, in the clinical consultation (Levine *et al*, 1994).

Clinical studies of risk aim to estimate both the size of any association and the degree of uncertainty in the central estimate. In the case of very rare events such as suicide, observational studies may be needed because randomised trials may simply lack sufficient power, even combined in meta-analysis. The basic analysis, however, is the same – a comparison of the occurrence of the event in patients exposed to the treatment with the occurrence of the event in a control group who do not receive the treatment. Various metrics are used in these studies, commonly the odds ratio, the risk ratio and the hazard ratio. Each of these has its own properties but all are essentially providing an estimate of risk of the event in the treated group compared with that in the control group. On the basis of the amount of statistical information in the study, a confidence interval can then be constructed around the risk estimate; this interval describes the range of values within which the true value lies.

An example

Using the data from Wheadon *et al* presented by Professor Healy, the relative risk of suicide in patients with bulimia treated with fluoxetine was 1.5 (95% CI 0.3–6.9). As Healy states, ‘with this confidence interval the data are potentially consistent with that risk being 6.9 times greater’. Actually, however, a more objective interpretation would be that the data are consistent with the risk being about 7 times greater or 70% less, with the most likely estimate being a 50% increase in the risk. This is clearly a very wide range of possible effects. Such a result is indeterminate – the risk could either be greatly increased or greatly reduced. However, although a potential benefit would be useful, the central estimate is of an increased risk and this would clearly be of clinical concern. The clinical implications of a 50% increase in the relative risk of suicidal acts would largely depend on the absolute risks.

In general, relative risks seem more impressive than absolute risks (Fahey *et al*, 1995). If, for example, the absolute rate of suicidal acts is high, say 30%, in the control group, then this would imply a 15% absolute increase to 45% in the treatment group, clearly a clinically significant difference. On the other hand, if the absolute risk was low, say 2%, then this would mean a rate of 3% in the treatment group – a much lower absolute difference. The size of the absolute risk of harm needs to be considered alongside the chances of benefiting from the therapy. A simple way of doing this quickly in the clinical consultation is to calculate the likelihood of being helped or harmed – a useful decision tool in which absolute risk is combined with patient-derived utilities (a way of rating preferences) of both helpful and harmful outcomes (Straus, 2002).

It is clear that even simple statistics can be presented in ways that encourage one particular interpretation or another. The task of the clinician is to provide the patient with an objective summary of the best available evidence, which the patient in turn can integrate with their own values and preferences in reaching a decision.

Declaration of interest

J.G. has received research funding and support from GlaxoSmithKline, Sanofi-Aventis, the UK Department of Health, the Medical Research Council and the Stanley Medical Research Institute.

References

- Altman, D. G. & Bland, J. M. (1995) Absence of evidence is not evidence of absence. *BMJ*, **311**, 485.
- Fahey, T., Griffiths, S., & Peters, T. J. (1995) Evidence based purchasing: understanding results of clinical trials and systematic reviews. *BMJ*, **311**, 1056–1059.
- Healy, D. (2006) The antidepressant tale: figures signifying nothing? *Advances in Psychiatric Treatment*, **12**, 320–327.
- Levine, M., Walter, S., Lee, H., *et al* (1994) Users' guides to the medical literature. IV: How to use an article about harm. *JAMA*, **271**, 1615–1619.
- Sackett, D. L., Rosenberg, W. M., Gray, J. A., *et al* (1996) Evidence based medicine: what it is and what it isn't. *BMJ*, **312**, 71–72.
- Straus, S. E. (2002) Individualizing treatment decisions: the likelihood of being helped or harmed. *Evaluation and the Health Professions*, **25**, 210–224.

John Geddes Professor of Epidemiological Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, UK. Email: john.geddes@psych.ox.ac.uk