



## Statistics notes: Absence of evidence is not evidence of absence

Douglas G Altman and J Martin Bland

*BMJ* 1995;311:485-

---

Updated information and services can be found at:  
<http://bmj.com/cgi/content/full/311/7003/485>

---

*These include:*

### References

This article cites 2 articles, 1 of which can be accessed free at:  
<http://bmj.com/cgi/content/full/311/7003/485#BIBL>

56 online articles that cite this article can be accessed at:  
<http://bmj.com/cgi/content/full/311/7003/485#otherarticles>

### Rapid responses

You can respond to this article at:  
<http://bmj.com/cgi/eletter-submit/311/7003/485>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top left of the article

---

### Notes

---

To order reprints follow the "Request Permissions" link in the navigation box

To subscribe to *BMJ* go to:  
<http://resources.bmj.com/bmj/subscribers>

- 13 Chen RT, Weierbach R, Bisoffi Z, Cutts F, Rhodes P, Ramarosan S, et al. A "post-honeymoon period" measles outbreak in Musinga Sector, Burundi. *Int J Epidemiol* 1994;23:1309-1314. Downloaded from [bmj.com](http://bmj.com) on 14 July 2009
- 14 Holt EA, Boulos R, Halsey NA, Boulos IM, Boulos C. Childhood survival in Haiti: Protective effect of measles vaccination. *Pediatrics* 1990;85:188-94.
- 15 Koenig MA, Khan MA, Wojtyniak B, Clemens JD, Chakraborty J, Fauveau V, et al. The impact of measles vaccination upon childhood mortality in Matlab, Bangladesh. *Bull World Health Organ* 1990;68:441-7.
- 16 Velema JP, Alihonou EM, Gandaho T, Hounye FH. Childhood mortality among users and non-users of primary health care in a rural West African community. *Int J Epidemiol* 1991;20:474-9.
- 17 Fleiss JL. The statistical basis of meta-analysis. *Stat Methods Med Res* 1993;2:121-45.
- 18 Garenne M, Aaby P. Pattern of exposure and measles mortality in Senegal. *J Inf Dis* 1990;161:1088-94.
- 19 Desgrées du Lou A, Pison G, Aaby P. The role of immunizations in the recent decline in childhood mortality and the changes in the female/male mortality ratio in rural Senegal. *Am J Epidemiol* (in press).
- 20 Harris MF. The safety of measles vaccine in severe illness. *S Afr Med J* 1979;38.
- 21 Aaby P, Bukh J, Hoff G, Leerhøj J, Lisse IM, Mordhorst CH, Pedersen IR. High measles mortality in infancy related to intensity of exposure. *J Pediatr* 1990;117:409-40.
- 22 Koenig MA. Mortality reductions from measles and tetanus immunization: a review of the evidence. In: Hill K, ed. *Child survival priorities in the 1990s*. Baltimore: Johns Hopkins Institute for International Programs, 1992:43-72.
- 23 Hartfield J, Morley D. Efficacy of measles vaccine. *Journal of Hygiene (Cambridge)* 1963;61:143-7.
- 24 Hull HF, Williams PJ, Oldfield F. Measles mortality and vaccine efficacy in rural West Africa. *Lancet* 1983;ii:972-5.
- 25 Griffin DE, Ward BJ, Esolen LM. Pathogenesis of measles virus infection: An hypothesis for altered responses. *J Infect Dis* 1994;170:S24-31.
- 26 Petralli JK, Merigan TC, Wilbur JR. Action of endogenous interferon against vaccinia infection in children. *Lancet* 1965;ii:401-5.
- 27 Rooth I, Sinani HM, Smedman L, Björkman A. A study of malaria infection during the acute stage of measles infection. *J Trop Med Hyg* 1991;94:195-8.
- 28 Gellin BG, Katz SL. Measles: state of the art and future directions. *J Infect Dis* 1994;170:S3-14.

(Accepted 15 June 1995)

## Statistics Notes

### Absence of evidence is not evidence of absence

Douglas G Altman, J Martin Bland

The non-equivalence of statistical significance and clinical importance has long been recognised, but this error of interpretation remains common. Although a significant result in a large study may sometimes not be clinically important, a far greater problem arises from misinterpretation of non-significant findings. By convention a P value greater than 5% ( $P > 0.05$ ) is called "not significant." Randomised controlled clinical trials that do not show a significant difference between the treatments being compared are often called "negative." This term wrongly implies that the study has shown that there is no difference, whereas usually all that has been shown is an absence of evidence of a difference. These are quite different statements.

The sample size of controlled trials is generally inadequate, with a consequent lack of power to detect real, and clinically worthwhile, differences in treatment. Freiman *et al*<sup>1</sup> found that only 30% of a sample of 71 trials published in the *New England Journal of Medicine* in 1978-9 with  $P > 0.1$  were large enough to have a 90% chance of detecting even a 50% difference in the effectiveness of the treatments being compared, and they found no improvement in a similar sample of trials published in 1988. To interpret all these "negative" trials as providing evidence of the ineffectiveness of new treatments is clearly wrong and foolhardy. The term "negative" should not be used in this context.<sup>2</sup>

A recent example is given by a trial comparing octreotide and sclerotherapy in patients with variceal bleeding.<sup>3</sup> The study was carried out on a sample of only 100 despite a reported calculation that suggested that 1800 patients were needed. This trial had only a 5% chance of getting a statistically significant result if the stated clinically worthwhile treatment difference truly existed. One consequence of such low statistical power was a wide confidence interval for the treatment difference. The authors concluded that the two treatments were equally effective despite a 95% confidence interval that included differences between the cure rates of the two treatments of up to 20 percentage points.

Similar evidence of the dangers of misinterpretation of non-significant results is found in numerous meta-analyses (overviews) of published trials, when few or none of the individual trials were statistically large enough. A dramatic example is provided by the overview of clinical trials evaluating fibrinolytic treatment (mostly streptokinase) for preventing

reinfarction after acute myocardial infarction. The overview of randomised controlled trials found a modest but clinically worthwhile (and highly significant) reduction in mortality of 22%,<sup>4</sup> but only five of the 24 trials had shown a statistically significant effect with  $P < 0.05$ . The lack of statistical significance of most of the individual trials led to a long delay before the true value of streptokinase was appreciated.

While it is usually reasonable not to accept a new treatment unless there is positive evidence in its favour, when issues of public health are concerned we must question whether the absence of evidence is a valid enough justification for inaction. A recent published example is the suggested link between some sudden infant deaths and antimony in cot mattresses. Statements about the absence of evidence are common—for example, in relation to the possible link between violent behaviour and exposure to violence on television and video, the possible harmful effects of pesticide residues in drinking water, the possible link between electromagnetic fields and leukaemia, and the possible transmission of bovine spongiform encephalopathy from cows. Can we be comfortable that the absence of clear evidence in such cases means that there is no risk or only a negligible one?

When we are told that "there is no evidence that A causes B" we should first ask whether absence of evidence means simply that there is no information at all. If there are data we should look for quantification of the association rather than just a P value. Where risks are small P values may well mislead: confidence intervals are likely to be wide, indicating considerable uncertainty. While we can never prove the absence of a relation, when necessary we should seek evidence against the link between A and B—for example, from case-control studies. The importance of carrying out such studies will relate to the seriousness of the postulated effect and how widespread is the exposure in the population.

1 Freiman JA, Chalmers TC, Smith H, Kuebler RR. The importance of beta, the type II error, and sample size in the design and interpretation of the randomized controlled trial: survey of two sets of "negative" trials. In: Bailar JC, Mosteller F, eds. *Medical uses of statistics*. 2nd ed. Boston, MA: NEJM Books, 1992:357-73.

2 Chalmers I. Proposal to outlaw the term "negative trial." *BMJ* 1985;290:1002.

3 Sung JY, Chung SC, Lai C-W, Chan FKL, Leung JWC, Yung M-L, Kassianides C, et al. Octreotide infusion or emergency sclerotherapy for variceal haemorrhage. *Lancet* 1993;342:637-41.

4 Yusuf S, Collins R, Peto R, Furberg C, Stampfer MJ, Goldhaber SZ, et al. Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: overview of results on mortality, reinfarction and side-effects from 33 randomized controlled trials. *Eur Heart J* 1985;6:556-85.

Medical Statistics  
Laboratory, Imperial  
Cancer Research Fund,  
London WC2A 3PX  
Douglas G Altman, head

Department of Public  
Health Sciences,  
St George's Hospital  
Medical School,  
London SW17 0RE  
J Martin Bland, reader in  
medical statistics

Correspondence to:  
Mr Altman.

BMJ 1995;311:485