Commentary: Statistics and death from meningococcal disease in children
Rafael Perera

Medical statisticians seldom directly make life and death decisions. Though I wouldn't like to have direct responsibility for making the decision to give a penicillin injection to a child with a purpuric rash in the community, I am conscious of the effect that my work may have on clinical decisions for such children. I feel a heavy responsibility when I conducted the statistical analysis of this paper.

In most datasets that I analyse, the main issue is to quantify if the effect observed is due to chance. The question is rarely about the direction of the main effect; instead it is more about the size of the effect and the precision with which it has been estimated. But on this occasion one key statistical decision determined the best estimate of the effect of parenteral penicillin given before admission to hospital was a modest benefit or substantial harm (table), and the statistician involved in the previous paper from the United Kingdom (on which current clinical policy is based) had taken the opposite view.

As the table shows, analysis A included all children with meningococcal disease for whom data were available and estimated a small protective effect of penicillin. Analysis B, which included only children in whom the general practitioner had suspected a diagnosis of meningococcal disease, estimated a substantial (six-fold) reduction in survival.

Simpson's paradox

This rather frightening statistical effect—actually changing the direction of the estimated effect from benefit to harm—is related to Simpson's paradox (or the Yule-Simpson effect), first described in 1952.

Simpson reported the seemingly impossible situation where the success of several subgroups seems to be reversed when the groups are combined.
The reason for the paradox is the combination of two factors: an imbalance in the proportion of each subgroup receiving each intervention and a different event rate in each subgroup. This was the case with the penicillin data. To have a chance of being given penicillin, children had to be seen by a general practitioner who suspected meningococcal disease, and children who were seen by a general practitioner had a lower mortality (18%) than those that were not (37%).

Analysis A was based on all the children with meningococcal disease in our study. It replicated previous work and was therefore reassuring. But on reflection and discussion with the clinicians, I realised it transgressed one basic statistical principle—it included in the analysis children who had no chance of receiving penicillin before admission. I therefore excluded the children who had not been seen by a general practitioner or in whom he or she had not diagnosed meningococcal disease (analysis B in the table). This analysis produced the evidence of substantially increased mortality.

Defining the population highlighted two important sources of confounding: the fast progression of the disease and the lack of specific signs and symptoms early in the illness. The analysis reported is based on a population composed only of children with a more slowly progressive disease (who had time to see their doctor) and in whom the signs and symptoms were specific enough for a diagnosis. The 158 children in whom the general practitioner diagnosed meningococcal disease were at a later stage of their illness than the 166 who also saw their general practitioner but were not so diagnosed (median time from onset of illness to consultation 14 v 8 hours). Furthermore, if the critical decision to administer penicillin in the 158 children is associated with severity of disease at the time (for example, more ill, higher chance that penicillin will be given) then the effect would be biased in the direction of penicillin causing harm. I thought it essential to adjust for severity of disease at the point at which the decision to give penicillin had been made.

Unfortunately, the limited data available made this difficult. The only validated measure of severity at diagnosis, GMSPS score, was assessed at admission to hospital—by which time penicillin is likely to have had an effect. Though severity scores at the time of diagnosis from the general practitioner’s notes were obtained, recording was incomplete. Nevertheless I used this partial assessment of severity at diagnosis, together with other recorded variables that are believed to be associated with mortality (such as type of disease), to obtain an adjusted effect of penicillin on mortality. Having adjusted with these variables I would have expected the association between penicillin and mortality to get weaker or disappear. The estimate adjusted for severity, however, showed a further increase in the association between penicillin and mortality (adjusted odds ratio 7.45, 95% confidence interval 1.47 to 37.67). The question still in my mind is whether the variables used did truly adjust for severity of disease.

I decided to write this commentary to highlight the major impact that simple statistical decisions can have on the results of clinical research; to increase awareness of the possibility of Simpson’s paradox, particularly in observational data of this nature; and to emphasise the importance of not assuming that strong associations are necessarily causal.

I thank Anthony Harnden, Richard Mayon-White, Matthew Thompson, and David Mant for help in preparing the manuscript.

Competing interests: None declared.


(Accepted 18 April 2006)

Corrections and clarifications

Cannabis intoxication and fatal road crashes in France: population based case-control study

The authors of this paper, Bernard Laumon and colleagues (BMJ 2006;331:1371-4), have alerted us to some wrong values in two of the tables in the full version (see bmj.com). In table 1 the correct values for Friday and Saturday night respectively are: all drivers, 532 (5.4%) and 649 (6.6%); positive for cannabis, 67 (9.8%) and 92 (13.5%); and positive for alcohol, 259 (12.4%) and 355 (16.5%). In table 4 the correct values for Friday and Saturday night respectively are: cases, 410 (6.1%) and 525 (7.7%); controls, 122 (4.1%) and 126 (4.2%). The authors confirm that these changes do not affect the conclusions of the paper.

Editor's choice: Improving on improvement

In her Editor’s choice of 6 May (BMJ 2006;332), Fiona Godlee mentioned Frank Davidoff’s draft publication guidelines for quality improvement reports. However, she mistakenly named his coauthor as David Batalden, whereas his real name is Paul Batalden.

Active and passive smoking and development of glucose intolerance among young adults in a prospective cohort: CARDIA study

In the methods section of this paper by Houston and colleagues (BMJ 2006;332:1064, 6 May), the definition of never smokers with passive smoke exposure should include all never smokers who reported having had passive smoke exposure or (not “and” as mistakenly stated) who had detectable serum cotinine concentrations (1-15 ng/ml).