Commentary: That was then, this is now

Nick Black

In January 2001, a national newspaper in the UK, the Sunday Times, published a supplement purporting to demonstrate the relative performances of British hospitals in terms of in-hospital mortality.¹ The data were supplied and promoted by a group of researchers from a leading medical school, internationally renowned for its scientific excellence. The publication attracted much public, political and media attention. Hospitals at the ‘top of the league’ congratulated themselves, while staff working in ‘badly performing’ hospitals were either demoralized, angry or, more sensibly, dismissed the findings. Here was another example of the misuse of inadequate observational data and yet further ammunition for the critics of such an approach. If only the researchers involved had read the report by FJ Double published 165 years earlier!²

That work elegantly highlights the potential shortcomings of using observational data to make meaningful comparisons of clinical effectiveness.² While Civiale’s claims regarding the aetiology of bladder calculi are measured and show awareness of the limitations of drawing conclusions from selected case series, his interpretation of the relative merits of lithotomy and lithotripsy is highly suspect. This surprises me, given his acknowledgement of the difference between a real and an apparent increase in the incidence of calculi (the latter arising from differences in the judgement and practice of individual clinicians) and the influence of fashion on clinical practice. He also demonstrates his awareness of the vagaries of basing incidence on surgical rates rather than prospective surveys of representative populations. Yet when it comes to comparing surgical techniques, he seems content to use crude post-operative mortality rates (20% versus 2.3%) even though he acknowledges the age mix of the patient populations differed considerably (50% under 14 years compared with 1%).³

In contrast, the report by Double and his colleagues recognizes that there are ‘numerous sources of error that are so difficult to avoid’. They, rightly, identify the potential confounding factors that might influence the sort of crude comparison carried out by Civiale: seasons, surgical difficulty, inflammation, bilious complications, illness duration, bladder damage and the general constitution of the patient. Not unreasonably, they conclude that the application of numerical methods in medicine is inevitably severely limited and that clinicians should continue to rely on intuition, experience and wisdom in deciding how to treat individual patients. Not unreasonable in 1835, but is such a conclusion reasonable in 2001?

Despite continuing examples of misleading use of observational data, such as that highlighted above, excellent examples do exist that demonstrate the potential application of these techniques.³ These can only be achieved if the data are accurate (valid and reliable) and complete, and if sufficient steps have been taken to adjust for case-mix or risk differences. Some people believe that the latter is never possible to achieve.⁵ If that view is accepted, much of health care will never be evaluated.

A more pragmatic view seeks a role for research in improving the quality of health care. This inevitably involves the use of observational data which, if conducted carefully, can make a major contribution.⁵ To take two recent examples, the demonstration of the danger of premature discharge from intensive care units at night⁶ and the relative merits of surgical procedures to correct stress incontinence in women.⁷

So, while the Parisian Academy was correct in 1835 to treat unadjusted crude comparisons based on selected case series with scepticism, 165 years later we have the information technology to allow us to collect high quality clinical data and the statistical techniques to make meaningful comparisons. The task is to ensure methodological rigour is achieved and that poor analyses do not damage further the reputation of observational approaches.

References