Commentary: How to derive causes from correlations in educational studies

I C McManus

Department of Psychology, University College London, London WC1E 6BT

I C McManus
professor of psychology and medical education

Ferguson and colleagues' paper contains much of methodological and substantive interest. Many *BMJ* readers will be unaware of structural equation modelling, a widespread technique in social research, the synonyms for which include path analysis, covariance modelling, latent variable modelling, and causal modelling. Structural equation modelling uses programs such as LISREL, EQS, and AMOS to fit models that combine multiple regression, factor analysis, psychometrics, and multigroup modelling and which answer subtle statistical questions.[1]

Hermann Goering famously (but erroneously) reached for his revolver on hearing the word culture. Causal modelling may have a similar effect in students taught in elementary statistics classes that "correlation does not imply causation." It doesn't, but that doesn't mean statisticians don't infer causation. They can, and they do, for science is about understanding causes.[2]

The problem of inferring causation is that if A and B correlate, then this may be because A causes B, B causes A, or that something else, X, causes both A and B. Although often presented as an intractable problem, it is far from that. David Hume in his *Treatise of Human Nature* of 1739, described the principle of priority of time whereby cause comes before effect. In the present case, the correlation between A level scores and preclinical performance cannot reasonably be interpreted as preclinical performance causing A level scores, since that would mean performing well at university caused students earlier in their lives to achieve better A level grades, which is nonsense. And so we infer A level scores cause preclinical performance.[3] The third option, that some third factor (X) causes both A level scores and preclinical performance, is directly testable if X has been measured (and so A level scores causing preclinical performance cannot result from both correlating with conscientiousness). If X has not been measured then the claim is not falsifiable, but structural equation modelling does help design the study which would make it testable.

Hume talked also of distant objects being "link'd by a chain of causes," the principle of contiguity. That can be seen in the academic backbone of this model. A level scores cause better performance indirectly, by causing better preclinical performance, which in turn causes better BMedSci performance and better clinical performance. A similar chain can be seen in a drinking song by Henry Purcell:

'Tis women make us love,
'Tis loving makes us sad,
'Tis sadness makes us drink,
And drinking makes us mad.

Ferguson and colleagues' paper has two important educational messages. Preclinical performance is predicted by conscientiousness, one of the "big five" personality dimensions, which meta-analysis confirms is often a predictor of job performance and job trainability.[4 5] Individuals with high conscientiousness see themselves as practical, thorough, and hardworking, rather than disorganised, lazy, and careless, and not surprisingly such individuals do better in preclinical examinations. Less obvious is the diminishing impact of conscientiousness on later performance, particularly clinical performance, perhaps because conscientiousness is less important for the self directed, more conceptual, less fact dominated learning required of clinical students. That is problematic for those wishing to use conscientiousness as a basis for student selection. Conscientiousness may be predictive of job performance only for repetitive, well organised, relatively closed tasks, and not for the more imaginative, thoughtful, open thinking required of an actor, an artist, a research scientist, or a creative clinician.

The other important result concerns the personal statement on the UCAS form, which although often claimed to have little validity, may be predictive but only for clinical performance. Once again, what is good for preclinical may not be good for clinical. Studies of student selection have to consider long term outcomes, not just the first one or two undergraduate years.

Competing interests: None declared.

1 Maruyama GM. *Basics of structural equation modelling*. Thousand Oaks, CA: Sage, 1998.
2 Pearl J. *Causality: models, reasoning, and inference*. Cambridge: Cambridge University Press, 2000.
3 Davis JA. *The logic of causal order*. London: Sage, 1985.
4 Matthews G, Deary IJ. *Personality traits*. Cambridge: Cambridge University Press, 1998.
5 Schmidt FL, Hunter JE. The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychol Bull* 1998;124:262-74.

Corrections and clarifications

Drug points
In their drug point reporting that leflunomide can potentiate the anticoagulant effect of warfarin, V Lim and I Pande stated in the last paragraph that the Committee on Safety of Medicines (CSM) had received over 300 reports of raised international normalised ratio (INR) in patients taking leflunomide concurrently with warfarin (*BMJ* 2002;325:1333). This is wrong. The CSM has in fact received over 300 reports of raised INR associated with warfarin and some other drug; four of these reports (up to mid-December 2002) were of raised INR associated with leflunomide. The authors say that the message remains the same, however—that care is needed when prescribing leflunomide in patients already taking warfarin.

Application of Framingham risk estimates to ethnic minorities in United Kingdom and implications for primary prevention of heart disease in general practice: cross sectional population based study
The authors of this paper, Francesco P Cappuccio and colleagues (*BMJ* 2002;325:1271-4), would like to thank Peter Macfarlane and his group in Glasgow for electronically coding about 1600 electrocardiograms according to the Minnesota rules.