EDITORIAL

Epidemiological methods to tackle causal questions

Michael Rutter

It is a dull day when there is not at least one media report of a claim that research has identified some new environmental cause of disease. Such claimed causes concern a wide range of supposed hazards including medical interventions such as the measles-mumps-rubella (MMR) vaccine, the thimerosal (mercury) preservative in other vaccines, dietary factors of many different kinds (coffee, alcohol, food additives, etc.), prenatal stresses or the intra-uterine exposure to the effects of maternal smoking or ingestion of alcohol, use of mobile phones, and living near radiation sources—to mention just a few examples. It is problematic, however, that many of these claims are not confirmed by other research and some are even reversed. Unsurprisingly both professionals and the lay public have developed a scepticism about claims on environmental causes of disease. Because much of the evidence derives from epidemiological studies of one kind or another, epidemiological science itself has come under scrutiny both within and outside the profession. Some have argued that only laboratory experiments and randomized controlled trials (RCT) can provide acceptable evidence on causation. That cannot be a solution, however, because so many of the putative environmental causes are ones that cannot be manipulated in humans—for a mixture of both ethical and practical reasons.

Because the identification of environmental causes is of such importance, and because the solution of the causal inference problem has no obvious single, simple solution, the Academy of Medical Sciences set up a working party (made up of clinical scientists, epidemiologists, statisticians, policy leaders and scientific media specialists) to answer the key questions of ‘how should we decide what to believe and when to take action?’ The report emphasized the need to focus on individual components in the causal process and not on some misleading abstract notion of a single basic cause. It also notes the many reasons why an observed association or correlation might not reflect an environmentally mediated causal effect. Such reasons include genetic mediation of the causal risk effect, social selection (allocation bias), and reverse causation to mention but three out of a much longer list of possibilities. It was also noted that the main problem in moving from an observed association to a causal inference did not lie in the effect of known, measured confounders, but rather in the effect of unknown, unmeasured confounders.

Statistical innovations

In recent years, there have been important statistical developments that go beyond the traditional multivariate regression techniques that seek to ‘control’ for confounders. Four approaches may be used as exemplars of the main strategies that are available. First, growth curve trajectories have been used to classify individuals according to their propensity or liability for some outcome of interest. Their value for the testing of postulated causal inferences is that they can be used to determine whether the occurrence of some particular environmental circumstance significantly alters the identified trajectory. Thus, Nagin et al. described such an application to test whether an individual’s participation in a delinquent peer group increased the likelihood of antisocial behaviour. As they pointed out, one advantage of this technique is that it also identifies whether any causal effect mainly applies to those with a high propensity for antisocial behaviour, or to those with a low propensity. The possible ‘confound’ being dealt with here concerns the overall liability to the disease/disorder outcome.

A rather different approach is provided by propensity scores. These aim to deal with social selection

MRC SGDP Centre, PO 80, Institute of Psychiatry, King’s College London, De Crespigny Park, Denmark Hill, London, SE5 8AF, UK. E-mail: Camilla.azis@iop.kcl.ac.uk
or allocation biases. The scores reflect the conditional probability of being exposed to the postulated causal agent, given the relevant background factors. One important value of propensity scores is that they identify levels of exposure that are so different in the groups to be compared that the key cells with minimal overlap should be omitted. A further use, however, is provided by the statistical techniques of inverse probability of treatment weighting (IPTW) devised by Robins. Expressed simply, the technique in effect creates two groups matched on their likelihood of risk exposure. Sampson et al. for example, used the approach to good effect to test whether marriage was associated with a within-individual reduction in antisocial behaviour.

The third approach is exemplified by direct acyclic graphs—so-called DAGs. Robins has argued that the operation of confounders should always be considered in terms of background knowledge of how they were likely to operate. The graphs are designed to make the probable pathways explicit, indicating the assumptions involved and how they may be tested. These then lead on to the appropriate statistical modelling to rule in, or rule out, the various postulated and tested alternatives. As always, much necessarily depends on the reality of the models and the quality of the background information available.

The fourth approach involves the concept of mediation. The objective here is different in that it is not concerned with identifying an environmental cause as such, but rather with testing whether an observed association between a risk factor and disorder outcome is mediated by a postulated mechanism involving some third variable. The rationale is that if the association between the independent variable ‘A’ and the outcome ‘C’ is mediated by the postulated variable ‘B’, the association between ‘A’ and ‘C’ should drop to zero when ‘B’ is included in the model.

**Natural experiments**

Advances in design features to test causal inferences are illustrated by the use of natural experiments—meaning design elements that provide an approximation to experimental conditions—to provide leverage on the causal inference. Essentially, the designs ‘pull apart’ variables that ordinarily go together and, by so doing, enable key direct contrasts between alternative hypotheses without the necessity for statistical manipulations that have to rely on some form of modelling to determine what the hypothetical findings would have been if the pattern of variables (and the associations) had been different.

Natural experiments can be considered in terms of five broad groups that differ in terms of what they seek to achieve. First, there are genetically sensitive designs that focus particularly on the need to differentiate between genetic and environmental mediation of risk effects—a need that arises as a result of gene-environment correlations. Multivariate twin designs do this by including both across-twin and across-trait analyses. Discordant twin pairs do so by contrasting disease/disorder outcomes, within pairs. Those who did and did not experience the relevant postnatal causal factors. Assisted conception allows the comparison of outcomes following the risk experience (such as prenatal exposure to the effects of maternal smoking) according to whether or not the mother and child are genetically related. With egg donation or surrogacy, for example, they are not. Children of Twins (COT) designs do so by capitalizing on the fact that the offspring of monozygotic twins are social cousins but genetic half-siblings. Adoption/fostering designs separate genetic and environmental mediation by contrasting rearing by biological and by nonbiological parents. Migration strategies provide yet another relevant contrast by utilizing the fact that when people migrate they bring their genes with them but may radically alter their lifestyle or be exposed to very different environmental circumstances.

A second group of genetic and related strategies uses sampling, not to control for genetic mediation, but rather to avoid a biased or confounded association between the putative environmental cause and the disease outcome. Thus, Mendelian Randomization (MR) uses molecular genetic polymorphisms as an instrumental variable that has a strong influence on the independent risk variable being studied but no direct association with the disease outcome other than through the environmental pathway leading from the independent risk variable. It was first introduced to deal with the possibility of reverse causation. Early onset of puberty follows the same instrumental variable rationale, albeit one not due to an identified gene. Adoption designs may be informative when adoption involves a move from a severely depriving environment to a good rearing environment after adoption. Within-individual change related to the rapid and radical change in rearing circumstances constitute the natural experiment provided that certain other conditions are met.

The third group of natural experiments have the main aim of avoiding selection bias by focusing on the very special situation when, unusually, the environmental hazard is experienced by the whole population. The Dutch famine in World War II provides one such example. The removal of the MMR vaccine in Japan but not the rest of the world provides a different example.

The final group consists of regression discontinuity design (RD). These capitalize on a major selection bias, provided that it is under strict control. The key defining feature is that allocation for some planned intervention is according to a strict pre-determined cut-off, rather than by randomization. The intervention effect is shown by a difference.
between slopes rather than between means (as in an RCT). Provided that certain key criteria are met, the RD design provides an unbiased estimate of a causal effect.

Contrasting the case studies with relatively strong causal claims with those giving rise to probably misleading causal claims showed that the successes shared several common features. First, they either concerned a very large effect (as with smoking and lung cancer) or they applied to rare and unusual outcomes with distinctive features (as with the foetal alcohol syndrome). Second, detailed careful attention was paid to alternative noncausal explanations of how to test for their possible role. Third, all made use of multiple research designs, including actual experiments. Fourth, the causal inference was made in multiple populations that differed in their characteristics. Fifth, animal models and human experimental studies contributed invaluable findings on biological processes, which complemented those of the rigorous high quality epidemiological studies.

I conclude that it is possible to use epidemiological methods to tackle causal questions but only if key criteria are met. The findings also indicate the falsity of the dichotomy between basic and applied research, with the need for an iterative interplay between the two and always for a seeking of ways using multiple research strategies in a hypothesis-testing style that typifies experimental medicine.

Conflicts of interest: None declared.

References

1 Academy of Medical Sciences. Identifying the environmental causes of disease: How should we decide what to believe and when to take action? London: Academy of Medical Sciences, 2007.


