The Royal Society of Medicine

Editorial

Journal of the Royal Society of Medicine; 2015, Vol. 108(1) 4–6
DOI: 10.1177/0141076814562713

Association and causation in epidemiology – half a century since the publication of Bradford Hill's interpretational guidance

Richard Wakeford
Centre for Occupational and Environmental Health, Institute of Population Health, The University of Manchester, Manchester, UK
Corresponding author: Richard Wakeford. Email: richard.wakeford@manchester.ac.uk

Austin Bradford Hill was one of the greats in the fields of epidemiology and medical statistics.[1] In the mid-20th century, with another great, Richard Doll, Bradford Hill initiated epidemiological studies that were to be highly influential in revealing the causal link between cigarette smoking and lung cancer.[2,3] However, this link was not accepted without a battle, and opponents of a direct cause-and-effect interpretation of the epidemiological association included such notables as the eminent statistician Ronald Fisher.[4] The debate spurred Bradford Hill to consider in some depth how the findings of epidemiological studies should be interpreted, and this led to the publication in 1965 in *Proceedings of the Royal Society of Medicine* of his seminal paper on association and causation.[5] To mark the 50th anniversary of the publication of this landmark paper, it is reproduced in this issue of *Journal of the Royal Society of Medicine*.

Bradford Hill's 1965 paper is a remarkable one that is full of insights. It proposes nine guidelines (often erroneously referred to as 'criteria', which Bradford Hill made clear they were not) against which a statistical association found in an epidemiological study may be judged as to whether a causal interpretation is reasonable or not. The most important of these guidelines are 'strength' (a strong association is more likely to be causal than a weak one), 'consistency' (an association is observed in different studies, under different circumstances, times and places), 'biological gradient' (i.e. dose-response – the effect should tend to be greater with a higher level of exposure) and 'temporality' (the effect follows the potential cause after an appropriate interval). Another guideline, 'biological plausibility', increases in importance as fundamental knowledge of disease aetiology accumulates, but such knowledge is clearly not complete, which should introduce a note of caution in placing too much emphasis on this guideline; nonetheless, the direct causal interpretation of certain statistical associations would stretch credibility given what is known today.

What is it about epidemiology that demands a detailed examination of the interpretation of its findings such as that conducted by Bradford Hill? With the exception of a few studies that are able to 'piggy-back' on clinical trials, epidemiology is an observational (i.e. non-experimental) science.[5–8] This is in contrast to randomised controlled clinical trials, which are experimental set-ups: study subjects are randomised between treatment groups by the investigators so that any background individual differences that might affect the outcome of the trial, even if unknown, are 'evened out' between the groups.[7] (Randomisation between treatment groups is a fundamental concept in clinical trials that Bradford Hill was key in establishing as a necessary requirement in study design.)[1,9] In observational epidemiology, randomisation of study subjects between groups (say, different levels of exposure to tobacco smoke) is not possible because studies in which people are deliberately exposed to potentially harmful substances without a realistic prospect of personal benefit are unethical. As a consequence, epidemiology must rely on data generated under the unconstrained conditions of everyday life, with no intervention on the part of the investigators, and this greatly complicates the interpretation of epidemiological findings.[7,8]

In clinical trials, essentially all that needs to be considered (assuming the study has been correctly designed, conducted and analysed) in judging causality is whether chance is a reasonable alternative explanation for the findings, i.e. how statistically significant are the results. In epidemiology, not only must statistical fluctuations be taken into account but also the potential presence of systematic errors (of primary concern is the existence of bias in study data, but other errors are possible) and confounding (when a factor considered in a study is associated with another factor that influences the outcome, producing a distorted, potentially misleading, result).[8]

Biases can be introduced into a study in many subtle (and not so subtle) ways. In case–control studies, for example, control selection bias is a notorious problem – controls differ from cases in ways that affect the results relating to the factor(s) under study.[10] Confounding as a credible explanation is always difficult (if not impossible) to entirely eliminate because even when potentially important confounding factors (such as smoking) are recognised and adjustment for their presence attempted, there remains the spectre of unknown extraneous influences lurking in the background. Thus, it is most unusual for an epidemiological study to generate findings that, in their own right, can be considered as unequivocally demonstrating a cause-and-effect relationship. An inherent mistrust of non-experimental studies is likely to have been a material reason underlying Ronald Fisher's questioning (initially, at least) of a direct causal interpretation of the association between cigarette smoking and lung cancer.[4]

Bradford Hill urges epidemiologists to carefully question the available evidence as to whether a causal interpretation of an association is reasonable or whether an alternative explanation is not just possible but perhaps probable. He summarises the situation thus:[5]

> None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a *sine qua non*. What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question – is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?
>
> No formal tests of significance can answer those questions. Such tests can, and should, remind us of the effects that the play of chance can create, and they will instruct us in the likely magnitude of those effects. Beyond that they contribute nothing to the 'proof' of our hypothesis.

What Bradford Hill produced was a framework for a structured approach to the interpretation of epidemiological findings. Note what he says of the role of statistical significance: highly significant results tell us that chance is unlikely to have produced the association (at least, not on its own) but that is far from demonstrating a cause-and-effect relationship. Indeed, an extremely significant result may indicate a causal link but could well have been generated by bias(es) in the study data (or another source of systematic error) and should act as a spur to examine the possibility of alternative explanations more intensely. It is quite understandable that epidemiologists want to find something important about disease causation and public or occupational health, but there is an obligation to try to stand back from a study and attempt to make an impartial assessment of the potential impact of bias and confounding on its results. This is not easy, since it is most unusual for observational data to be free of any source of bias – it is a question of whether investigators can dispassionately judge the degree of influence that might be attributed to bias; but in the end, this may be best left to independent review.

Half a century after the publication of Bradford Hill's detailed examination of epidemiological association and causation, his paper is still of substantial relevance today, possibly more so given the number of epidemiological studies that are now undertaken. Observational epidemiology has made major contributions to the establishment of causal links between exposures and disease and plays a crucial role in, for example, the evaluation of the International Agency for Research on Cancer of the carcinogenicity of a wide range of human exposures;[11] but the 'positive' findings of epidemiological studies have often proved to be controversial and transitory, leading to scepticism, if not cynicism.[12] Perhaps Bradford Hill's paper should be required reading, not just by epidemiologists but more generally by the medical profession, and beyond?

Declarations

Competing interests: None declared

Funding: None declared

Ethical approval: Not applicable

Guarantor: RW

Contributorship: Sole author

Acknowledgements: None

Provenance: Commissioned; editorial review

References

1. Doll R. Sir Austin Bradford Hill and the progress of medical science. *BMJ* 1992; 305: 1521–1526.
2. Doll R and Hill AB. Smoking and carcinoma of the lung. *BMJ* 1950; ii: 739–748.
3. Doll R and Hill AB. A study of the aetiology of carcinoma of the lung. *BMJ* 1952; ii: 1271–1286.
4. Doll R. Proof of causality: deduction from epidemiological observation. *Perspect Biol Med* 2002; 45: 499–515.
5. Hill AB. The environment and disease: association and causation. *Proc R Soc Med* 1965; 58: 295–300.
6. Hill AB and Hill ID. *Bradford Hill's Principles of Medical Statistics*. London: Edward Arnold, 1991.
7. Grimes DA and Schulz KF. An overview of clinical research: the lay of the land. *Lancet* 2002; 359: 57–61.

8. Grimes DA and Schulz KF. Bias and causal associations in observational research. *Lancet* 2002; 359: 248–252.
9. Schulz KF and Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. *Lancet* 2002; 359: 515–519.
10. Grimes DA and Schulz KF. Compared to what? Finding controls for case-control studies. *Lancet* 2005; 365: 1429–1433.
11. International Agency for Research on Cancer (IARC). *IARC Monographs on the Evaluation of Carcinogenic Risk to Humans: Preamble*. Lyon: IARC, 2006.
12. Taubes G. Epidemiology faces its limits. *Science* 1995; 269: 164–169.