Commentary: Balancing quantity and quality when designing epidemiological studies

Andrew Phillips

The increasing identification of genetic factors which predispose people to higher risk of certain diseases will reinvigorate the epidemiology of such diseases and provide new challenges for study design. Some of these design issues are addressed in the paper in this issue of the International Journal of Epidemiology by Wong et al., which provides a useful formula for calculating sample size estimates for detecting gene–environment interactions for continuous traits.1 An example of such a potential interaction—used in the paper—is where a study aims to see if the association between physical activity and insulin sensitivity (both continuous traits) differs between two genetically defined subgroups. The important additional factor considered by Wong et al.—too often neglected in study design considerations—is the degree of precision with which the exposure (in this case physical activity) and the outcome (insulin sensitivity) can be measured. For a given sample size, the more imprecisely an exposure and outcome are measured, the lower the power of the study to detect the interaction. Often there is flexibility in the choice of measures of exposures and outcomes. Measurements which can be feasibly performed on very large numbers of people may be relatively imprecise compared with more expensive alternatives. A particular example of this flexibility arises in the situation where a substantial proportion of the measurement imprecision of a single exposure measurement is random. Here, precision can be improved by taking two or more exposure measurements per person. Obviously, for a fixed total number of subject evaluations, the use of two measurements per subject would result in a halving of sample size. So, as Wong et al. reiterate, ‘the practical consideration when designing studies … (is) the trade-off between sample size and measurement precision’.

So, which factor should weigh more heavily? This fundamental issue has been examined previously outside of the gene–environment interaction case2–4 and the same conclusion has been arrived at as that of Wong et al.; i.e. ‘our calculations suggest that this trade-off should be weighted towards better measurement’. How much can it matter? To continue Wong et al.’s physical activity/insulin resistance example, the choice could be between use of a questionnaire to characterize physical activity, or two repeated measurements of heart rate. The former is correlated with the true exposure of interest (habitual energy expenditure) with correlation only around 0.3,5 while for the latter the correlation is reported as around 0.88.5 Sample sizes needed to detect a difference in the physical activity/insulin resistance association between those in two gene allele groups would be something approaching 100 times smaller if the latter measurement were used, with fasting insulin as outcome, than if poorer measurements were used. Use of more precise measurements has the added advantage that, not only will it usually result in greater power to detect the existence of an association (including interactions), but there will also be less bias in the estimates of the magnitude of such associations.

There is, of course, a limit to how much smaller an epidemiological study can become in the interests of using better measurements, so cost will always be an issue. In some areas of research, expensive measurements have been made regularly (e.g. 3 monthly) on many thousands of individuals and these measurements related to outcomes of serious clinical disease and death. Such areas are that of human immunodeficiency virus (HIV) infection and, increasingly, hepatitis C virus (HCV) infection, where expensive measures of plasma viral load, costing around £70 per test, have been related to risk of development of AIDS7 or liver failure,8 respectively. A measure of the concentration of CD4+ lymphocytes in peripheral blood is also used in HIV, as a measure of immune competence against potential AIDS-defining opportunistic infections. For HIV, even 5 years of follow-up of a typical clinic population of 2000 people with standard monitoring tests costs over £5 million in laboratory tests alone. This has been possible due to the increasing integration of clinical medicine and epidemiology in these and some other infectious disease areas. In HIV, for example, we have a defined population of infected individuals at risk (of AIDS) that are clinically well. Such people are followed up at regular intervals and the aim of clinicians and epidemiologists together has been to precisely describe the relationship between such factors and risk of disease.7,9 This is so that their role in patient monitoring can be further developed, leading to better planning of treatment strategies such as when to start and change antiretroviral therapy.

With the sequencing of the human genome, such a situation may become more common in other disease areas. Some genetic factors identified may well prove to be necessary or almost necessary (though perhaps not sufficient) for a given disease to occur. Where a genetically distinct asymptomatic sub-population with a high lifetime risk of development of a serious disease can be identified this will potentially transform clinical medicine for that disease, so that there will be clinical follow-up—albeit very infrequent in some cases—of all those identified with the genetic predisposition. This could similarly result in the coming together of epidemiology and clinical medicine to address important questions relating to individuals’ prognosis and possible need for such interventions as are available.

However, as we all face the growing impact of genetics on our area of epidemiological interest, it is important that we remember old lessons as we attempt to adapt to the new situation. Perhaps the most important such lesson, that Wong et al.’s work reiterates and extends, is to be wary of compromising on measurement precision for the sake of obtaining a larger study.
References


