A high response rate has long been regarded as one of the hallmarks of a ‘good’ epidemiological study. However, the validity of results from studies that require participants to complete questionnaires, have an interview or provide a biological specimen is becoming increasingly jeopardized by decreasing participation. One analysis showed that cooperation rates (participants interviewed/number eligible that were contacted) in published population-based case–control studies declined by 3.33% per year in cases and 5.15% per year in controls from 1991 to 2003.1 In such studies, cooperation rates of 70% and response rates (participants interviewed/(participants interviewed + eligible non-participants + people of presumed but unconfirmed eligibility) of 50% are not uncommon today.2 Furthermore, and maybe because of declining rates, only a minority of studies published sufficient information for either cooperation or response rates to be calculated.1

A fundamental tenet of a case–control study design is that the controls should be selected randomly, and independently of exposure status, from the source population from which the cases were drawn and thus be representative of the exposure experience of the source population.3 Poor response rates by themselves do not create selection bias any more than high response rates guarantee unbiased estimates, but by endangering representativeness, low participation can provide more opportunity for bias to occur. Selection bias develops if the selection probabilities are different for cases and controls based on their exposure status.

Since odds ratios maybe biased due to low participation, should we discard case–control studies as a valid study design?

The advantages of case–control methods are well known. These include the ability to investigate causes of rare diseases, efficiency of time as they use historical information, relative affordability, statistical efficiency as they require relatively few subjects for sufficient power and are reasonably simple to analyse, and they allow the investigation of multiple putative disease causes.

In spite of the drawbacks of low and decreasing participation rates, it is important that we do not throw the baby out with the bath water. Given that case–control studies are extremely useful for examining causality in rare diseases, and given that response rates are unlikely to improve in the near future, case–control studies can contribute significantly to future epidemiologic research if bias can be adequately investigated and quantified.

Is there a viable alternative to using population controls in population-based case–control studies?

Some investigators have used other sources of controls, e.g. hospital patients, to increase response
rates, but these have been shown to often produce considerable selection bias. Milne et al.\(^5\) have investigated another method of addressing bias arising from poor response rates of population controls by using sister controls instead. They present a very interesting analysis of risk factors for breast cancer using the two sets of controls, but do the sister controls really ‘provide more valid risk estimates’?

If we just consider the Australian data in the original study from 1992 to 1995, the response rates were 73% for cases and 64% for controls.\(^5\) Using information from both the original study and the current study, the response rates for the new sample appeared to be 67% (1111/1659) for cases and 68% for both population (613/899) and sister (1390/2039) controls. However, the sister controls were ‘matched’ to only 53% (771/1465) of cases.

In addition to potential selection bias from low response rates in cases, this design introduces two potentially major sources of selection bias in the sister controls. First, it appears that only half the cases had eligible sisters from which to choose the controls, and, secondly, only 68% of these eligible sisters participated. So selecting sister controls did not ‘solve’ the problem of motivation to participate. Furthermore, given the potential selection bias from both the availability of and low response rates in sister controls, using these controls would likely make any effect estimates less valid than population controls.

**But are high response rates necessarily better?**

We have generally obtained high response rates in our population-based case-control studies in New Zealand: 88% in cases and 80% in controls in a 1996–98 study of prostate cancer,\(^6\) and 80% in cases and 74% in controls in a 2007 study of colorectal cancer.\(^7\) But maybe we have been too complacent and not examined potential selection bias sufficiently?

**So how should selection bias be assessed?**

A thorough assessment of the effects of selection bias should be carried out for every case-control study. If no such assessment is performed, then any discussion of bias degenerates into mere hand-waving.

In order to properly assess the influence of participation on validity at various stages of the study design, it is recommended that authors report both response and cooperation rates with a transparent explanation or flow chart of their calculation. In the interests of brevity, this could be included as an appendix or additional data online. Where possible, the demographics of non-responders and suitable external populations should also be included.

There is considerable literature concerning the many methods to minimize, or quantify the impact of, or analytically control for selection bias, depending on the availability of appropriate data. These techniques include careful study design and confounding adjustment, using multiple strategies to contact potential participants, employing incentives to improve willingness to participate, using external population data such as the census to compare demographic distributions between responders and non-responders in population-based studies, attempting to collect relevant data from a subsample of non-respondents, using sensitivity analyses that assume (plausible) best- and worst-case scenarios in the non-responders,\(^8\) carrying out simulation studies exploring the effects of non-response or by using probability weighted methods\(^9\) or multiple imputation\(^10\) as part of the statistical analysis.

To maximize the information from the Milne study,\(^4\) we encourage the authors to prepare a methodological paper using the Australian data, including a flow chart of participation; demographics (and other available data) of case responders and non-responders, population control non-responders, responding cases with and without eligible sisters and participating and non-participating sisters; comparison with census or other appropriate external data; and quantification of the bias using a statistical method and sensitivity analysis. This would be a valuable contribution to the discussion of alternative control selection and selection bias in case-control studies.

**Conflict of interest:** None declared.

**References**


