Epidemiological research

In the second article of our series on epidemiology, Mona Okasha explains the different types of study involved.

Epidemiology may seem difficult to understand. It has concepts and words very different to those that you come across in the rest of your medical training. In this article I give an overview of what types of studies are available to us, with examples of when and when not to consider each type. I have not included an exhaustive list of advantages and disadvantages; these are available in most textbooks. The next article will concentrate on how to interpret study results.

How do I choose a study type?

(1) Ensure that it directly addresses your research question. There is no point performing a study which will not answer the question in hand. Many study types can address the same question, so choose a design carefully.

(2) Make sure the study is ethical. Epidemiology usually involves the active cooperation of patients. They usually participate for the good of research rather than for any benefit which they stand to gain. The ideal study must be ethically sound and ought to minimise the invasiveness to the participants.

(3) Work within your budget. Epidemiological studies tend to be much larger than other scientific experiments and resources are always limited. Patient recruitment and assessment take time and money. You need to consider the cost of collection and testing of biological samples, telephone calls and postage, and of course salaries—personnel resources have high costs attached. The cost of the research can preclude certain study designs.

(4) Make sure the results will be valid. The most frequent flaws in epidemiology are bias and confounding. Bias produces incorrect results and can rarely be rectified after data collection. Confounding can be addressed during statistical analysis (number crunching), provided that sufficient information was gathered during the study. Ways to avoid these two hazards will be discussed in the next article.

Hierarchy of evidence

There are six main types of epidemiological study to choose from. Each design is ranked according to what was traditionally thought of as its usefulness in providing accurate results. This ranking is sometimes referred to as the hierarchy of evidence, with stronger evidence coming from the studies later rather than earlier on in the list.

Case series

Lowest in the pecking order is the case series. This is simply a description of cases. Importantly, there is no comparison group and so case series are often not considered as epidemiology. They describe patients’ characteristics, and may generate ideas for future studies. They can also be misleading though. For example, in the early 1980s, doctors noticed that patients with immune system deficiencies tended to be young homosexuals. That observation suggested that their symptoms were caused by the use of “poppers” (amyl nitrates—sexual stimulants). Experiments of the immunological properties of amyl nitrates were underway before the HIV virus was identified. The use of case series is clearly limited, although that should not deter clinicians from observing monitoring of unusual patients.

Ecological studies

Second in our hierarchy comes the ecological study, when we compare groups of people not individuals. Assuming that associations seen on a group level also hold on an individual level leads to ecological bias. Consider a study of suicide and religion. Higher suicide rates occurred in regions where a higher proportion of Protestants lived. To infer that Protestants are more likely to commit suicide than people of other religions may be incorrect. Although the rates are higher in the regions with more Protestants, we do not know the religion of the people committing suicide. Suicide could be more common among religious minorities who live in predominantly Protestant regions. However, with careful interpretation of results, ecological studies do have benefits. They are generally quick, cheap, and can be performed from data which are published routinely—for example, death

The six study designs that you need to know

- Case series—what clinicians see
- Ecological—geographical comparisons
- Cross sectional—survey, a snapshot in time
- Case-control—compare people with and without a disease
- Cohort—follow people over time to see who gets the disease
- Randomised controlled trial (RCT)—the human experiment
Considerations when choosing a study type

- Does this design address my question?
- Is this study design ethical?
- What resources do I have (time, money, personnel)?
- Is there a cheaper or quicker way of answering the same question?

Cross-sectional study

A cross-sectional study is a snapshot in time—that is, a survey. Participants are picked from a well-defined population. Exit polls on voting day are cross-sectional studies, where the population is all voters exiting from a particular polling station. We contact people only once, so these studies are relatively cheap. But that limits their usefulness, since we can study only current diseases (prevalence), and cannot identify when people first get a disease (incidence). If we are interested in whether a certain behaviour may cause a disease we may infer incorrect results. People may change their behaviour once they get a disease. This concept of reverse causality will be discussed in the next article.

Case-control studies

Choosing your participants on account of their disease is the basis of a case-control study. On face value, this seems a simple study design, where cases are people with the condition and controls are those without. Suitable control selection can be tricky, and unsuitable controls can invalidate your results. Suppose that you want to know whether smokers are at an increased risk of colon cancer. Comparisons are made between the smoking habits (probably current and past) of cases and controls. It is no good choosing controls who have lung cancer or heart disease, as they will be more likely to be smokers than the population from which the cases came, and that will distort your results. Another risk of case-control studies is recall bias, which occurs if cases and controls recall past events differently. Because we actively choose the cases, a case-control study ensures that we find enough people when a disease is rare. These studies are rarely suitable for investigating causes of death, since dead people cannot provide vital information. Carefully designed case-control studies can provide useful results, and such studies should not be ignored without a careful evaluation of the methodology used.

Cohort studies

In a cohort or follow-up study, a healthy group of people (cohort) are identified. They are then followed over time to see who develops the disease of interest and who does not. The important point to remember about cohort studies is the time factor—that at the beginning of the study neither the people themselves nor the researchers know who is going to get what disease. This effectively avoids recall bias, although other types of bias can still hinder these studies. However, the costs of cohort studies are sometimes prohibitive. If you are studying mobile phone use and cancer development you need to follow up people for a long time before the cancer becomes evident. The resource implications is often why cohort studies are not chosen in epidemiology.

Randomised controlled trials

Finally, what is often called the gold standard of epidemiological studies—the randomised controlled trial or clinical trial. This is a human experiment, where people are randomly assigned to receive one treatment or the other. This treatment is often a drug as clinical trials of drugs are required before being licensed for prescription. The treatment may also be a health intervention, such as a weight loss or smoking cessation program. It is necessary that the individual (and preferably also the health professional) are “blind” to which treatment a person receives, although this is not always possible. In clinical trials, the aim is to replicate the “real life” situation, so that the results obtained are as close to what would happen if the treatment was used in a real life situation. Clinical trials tend to be extremely expensive, and are unsuitable for use in some situations. For example, it is not ethical to randomise people to heavy drinking to investigate the effect of its use on breast cancer.

Example of study design

As an example of how to choose a study, I will return to the study of orgasms and mortality in middle aged men.1 Which of the study types described above would answer the question of whether sex and death are related?

A case series of people who have died would not be useful—given the frequency of the exposure (orgasms) and the outcome (death), there would be nothing that would stand out in a study of all people who died. Remember, there is no comparison group.

In an ecological study, we would compare rates of orgasm and death across different geographical areas. But how would we find out orgasm rates per area? And would these differ sufficiently across areas to be able to correlate them with death rates?

A cross-sectional study by its nature excludes dead people. We could do a survey of women, asking them about the frequency of their partner’s orgasms, and ask if the partner is still alive. But how accurate are reports from women regarding their partner’s orgasms? (That approach might be more accurate than asking men directly!) In addition, it may be too distressing to ask recently widowed women about their sex life.

Similar problems would be encountered with a case-control study. Such studies of mortality always rely on proxy information—for example, from a partner rather than the individual. The cases would be people who had recently died. But who would be the controls? Brothers? Neighbours?

The method the authors chose was the cohort study, when people were asked about their sex life then followed up for 10 years. This is not without its problems, although these are more issues of interpretation than design, so will be considered in the next article.

Finally, could they have done a randomised controlled trial, allocating men to groups that have frequent, infrequent, or no sex. In that case recruitment would have been pretty difficult—would you open yourself to the option of no sex?

And finally...

As you will have gathered from the series of questions above, there is no clear cut answer to which study design to use. Often two or more designs are possible, and then practical (“how much time and money do I have?”) and qualitative (“which study will give better results?”) considerations help to make the choice. The ability to choose a useful study design is an art which improves with practice.

Mona Okasha epidemiologist, University of Bristol
mona.okasha@bristol.ac.uk


Further reading