Interpreting epidemiological findings

In the third article in this series **Mona Okasha** gives a step by step guide to understanding an epidemiological study

Last month's article considered important features of epidemiological study design.[1] The focus of this article is how to interpret the study's results. This guide is equally applicable to a study of your own as to published journal articles. Using this same structure you will be able to evaluate an epidemiological study that you are faced with in an exam situation.

Between study design and interpreting results is a wide gap—the statistical analysis. Statistics are used to come up with the numbers which form the results. That is a whole subject area in itself, and I refer you to Kirkwood's book (see Further reading) for a clear overview of the subject.

Considerations for evaluating epidemiological results

- Chance—random variation.
- Bias—deviation from the truth; selection bias, when participants are not chosen at random; information bias, when the accuracy of data differs according to exposure or outcome.
- Confounding—an alternative explanation.
- Reverse causality—perhaps the outcome caused the exposure and not vice versa.
- Causal association—this is a very strong assumption, beware.

The language of epidemiology

As with other medical disciplines, epidemiology has its own vocabulary. I will use what may be to you unfamiliar words throughout this article. Explanations of these are given in the Glossary. Epidemiology is used to describe associations between exposures and outcomes. In the study of sex and death, which investigated whether middle aged men who had frequent orgasms were more or less likely to die than their less sexually active peers, the exposure is frequency of orgasm and the outcome is death.[2]

The results on face value

The results of epidemiological studies are expressed as a comparison of the outcome between two or more groups. In the sex study, we want to know whether the death rate is equal among groups of men who have differing frequencies of orgasms. We calculate the rate among men with infrequent and those with frequent orgasms. Divide the former by the latter to calculate the rate ratio (RR). If the rates in the two groups are equal, the RR is 1. RR $<$ 1 implies a lower rate among the men with few orgasms compared with frequent orgasms. RR $>$ 1 means the rate among the men with few orgasms is lower than those with no orgasms. In the sex study, the odds ratio (OR) was used to approximate the RR. The value calculated was 2.0, indicating that men with low frequency of orgasms were twice as likely to die in the follow up period than men with frequent orgasms.

Do we believe these results?

Statistics such as the RR are used to describe relationships. We then need to decide how much worth to put on these results. We are going to consider the results under five headings, (see box). The order in which they are listed is important, because if at any stage you decide that the study is not sound you need to bear this in mind for each of the subsequent considerations.

Chance

Statistics help us to decide whether our results may be chance findings. There is no definite cut-off point that you can use to say, "This was chance" or "That was not chance." That was done in the past, by the use of probability (P) values. That method is no longer considered sufficient in epidemiology.[3] Whether to believe a result is due to chance or not is a qualitative decision, assisted by calculating a confidence interval (CI). The OR is only an estimate. If you did the same study again you would probably get a slightly different result. The CI gives us a range within which we can be relatively sure that the "real" OR lies. In the sex study the 95% CI is from 1.1 to 3.5. Loosely interpreted, we can be 95% sure that the real OR lies between these values. Remember that if rates in the two groups are equal the OR would be 1.0.

Bias

If we are quite happy that our results were not simply a matter of chance we then consider whether they may be biased or simply wrong. Bias is generally a problem of study design or data collection, and ought to be avoided at the outset of the study. If you are interpreting someone else's results consider that bias may have occurred at every point of the study. There are two main sorts of bias—selection bias and information bias.

Selection bias

Selection bias is an error of which people were chosen to be in the study. For bias to occur, the reason that you chose to include or exclude people must be related to the exposure or outcome of interest. Selection bias is a particular issue in case control studies. Imagine a study of alcohol drinking and motor cycle crashes. We could choose our cases from an accident and emergency department. Controls from the same hospital would be poorer, smoke more, and drink more (common characteristics of an inpatient population) than the general population. Since the reason for choosing controls (being an inpatient) is related to our exposure (drinking alcohol) our results may be biased. Having carried out the study, there is very little that we can do to rectify the matter.

In cohort studies selection bias is an issue if people drop out of the study, for reasons which are related to the exposure or outcome. In the sex study men who did not participate tended to be older, shorter, and more likely to be in manual occupations than those who gave responses to the orgasm question at the outset. The men excluded may have been likely to have frequent orgasms and more likely to die. The ratio between the mortality rates that was found in this study is probably smaller than the ratio that would have been found if all men had been included. Selection bias can work in the opposite direction also, magnifying the results beyond the "true" value.

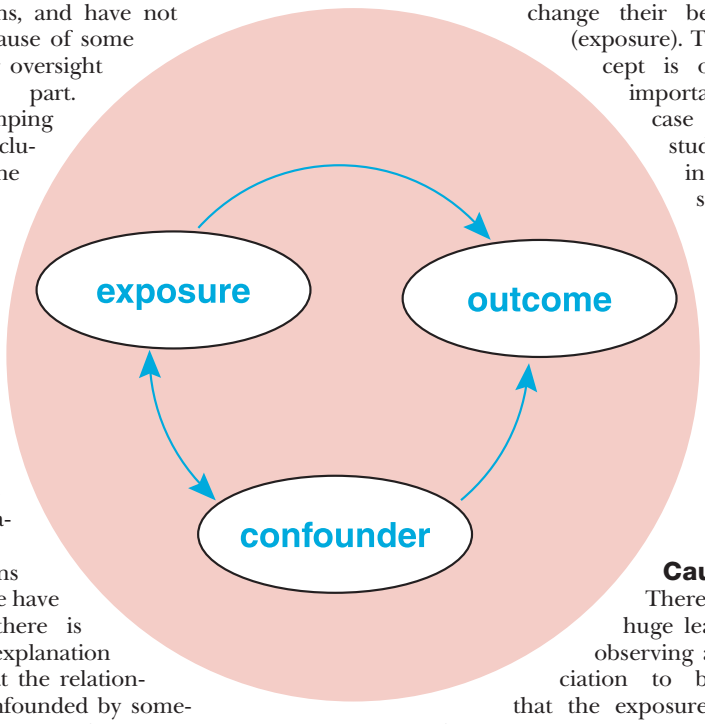
Information bias

Information bias occurs when we collect information wrongly. It is also known as differential measurement error, when we measure things differently between cases and non-cases or controls. In case control studies information bias occurs most frequently when cases and controls remember (or tell us) things differently. People with a disease may be more likely than controls to say that they were thinner or lighter smokers or some other socially desirable attribute. This is known as recall bias, and is an important source of error in case control studies. In cohort studies information bias is most common when measuring the outcome. For example, in studies of a treatment, people receiving

treatment are monitored more closely than "healthy" people. Discovering the outcome (case ascertainment) may be different between the exposed and non-exposed group. In the sex study there is no reason to think that the authors were more likely to find out about the deaths of men who had frequent or less frequent orgasms, so information bias is unlikely to be an issue.

Confounding

We have now got as far as thinking that our results are not just random observations, and have not arisen because of some mistake or oversight on our part. Before jumping to the conclusion that the exposure causes or prevents the outcome—for example, sex prevents death—we need to consider alternative explanations for what we have seen. If there is another explanation we say that the relationship is confounded by something. The "something" is called a confounder, or a confounding factor. To be a confounder, the factor needs to be related to the exposure and be a risk factor for the outcome (see figure).



The ubiquitous potential confounder is age. Most diseases increase in frequency with increasing age (so the confounder is a risk factor for the outcome). Many exposures, especially behavioural factors, are also related to age. Unlike bias, there are things that we can do to cope with confounding. Using statistics, we can adjust or control results for confounding factors. Adjustment takes into account the confounder and removes its effect. Comparing the unadjusted (crude) results with the adjusted results gives an indication of whether there is confounding present. In the sex study the odds ratio changed from 2.0 to 1.9 after adjusting for age, social class, systolic blood pressure, smoking, and coronary heart disease at baseline, suggesting no confounding. We can adjust only for factors which we know about though, and residual confounding may persist when the adjusted results still suffer from confounding by another variable.

Reverse causality

We have now reached the stage where we think that the results we have are not explained by chance, bias, or confounding. Before considering whether the exposure may cause the outcome we must consider whether the outcome may have caused the exposure. This concept is known as reverse causality. In the sex and death study that is a likely explanation of the results. Men who had a chronic illness at the beginning of the study, so were more likely to die over the follow up period, were less likely to have been sexually active. The disease (outcome) caused individuals to change their behaviour (exposure). This concept is of great importance in case control studies and in cohort studies with short follow up periods.

Causality

There is a huge leap from observing an association to believing that the exposure causes the outcome. Criteria which ought to be fulfilled before assuming causality were drawn up by Austin Bradford Hill in 1965.[4] Briefly, these criteria require a consistent body of evidence to have accumulated. The exposure must clearly predate the outcome and there ought to be a biological or other rational explanation of the results. It is important to remember that a single epidemiological study is never sufficient to determine a causal relationship.

So is sex good for you?

You now ought to be able to critically appraise Davey Smith et al's article, and make an informed decision as to whether you believe that a lack of frequent orgasms could increase your risk of dying. Bear in mind that there is never a right or wrong answer. Appraising epidemiology is about weighing the evidence and coming to a conclusion which is consistent with evidence from both experimental and observational work.

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

1 Okasha M. Epidemiological research. *studentBMJ* 2001;9:277-8.(August.)

2 Davey Smith G, Frankel S, Yarnell J. Sex and death: are they related? Findings from the Caerphilly cohort study. *BMJ* 1997;315:1641-4.

3 Sterne J, Davey Smith G. Sifting the evidence–what's wrong with significance tests? *BMJ* 2001;322:226-31.

4 Hill AB. The environment and disease: association or causation? *J R Soc Med* 1965;58:295-300.

Further reading

Kirkwood BR. *Essentials of medical statistics.* Oxford:Blackwell Science, 1988.
Last JM. *A dictionary of epidemiology.* Oxford: Oxford University Press, 1988.

Glossary

- Adjust—to alter results because of how a confounder affects the association.
- Association—the relationship between an exposure and an outcome.
- Case ascertainment—finding who has a particular disease, through screening, searching medical records, or asking individuals themselves.
- Confidence interval—a measure of how precise your results are. A narrow confidence interval implies precision, a wide confidence interval imprecision.
- Exposure—the factor we think may alter disease risk.
- Odds ratio—the ratio between the odds of disease in one group compared with another. Odds are used to approximate risk.
- Outcome—the disease or condition that we are interested in.
- Rate—the risk of getting a disease over a particular time period.
- Rate ratio—the ratio between the rate of disease in one group compared with another.
- Recall bias—inaccuracies occurring when cases and controls recall past events with different levels of accuracy.
- Residual confounding—confounding occurring in adjusted results, either because we have not controlled for a particular factor, or because a confounding factor was inaccurately measured, so adjustment has not taken away the whole effect of that confounder.
- Risk factor—something associated with an altered probability of disease, not necessarily causal.