**5**

UNIT 5 - Association

Causation, Effect Modification and Validity

# Introduction

In Unit 1 we introduced the concept of causality in epidemiology and presented different ways in which causes can be understood and classified. Then in Unit 3 introduced epidemiological study designs and in Unit 4 we further explored in more depth randomised controlled trails, Quasi-experimental studies and ecological studies. We now want to pause and put together what we covered in Unit 1 on general aspects of causation with the epidemiological principles we learnt in Unit 3 and the methodology used in conducting study designs which we explored in Unit 3. Building on those concepts, in this Unit we present the practical steps epidemiologists follow when answering the basic question underlying most of their work: **does exposure X cause outcome Y**?

Contents

*Session 1* explains under which conditions an observed association between two epidemiological variables indicates the existence of a causal relationship.

*Session 2* introduces complex causal relationships, in which more than two variables are involved.

Finally, the short *Session 3* discusses the concept of validity of epidemiological studies, and contrasts internal vs. external validity.

# Unit 5 - Session 1: Association and Causation

## Introduction

A principal aim of epidemiology is to assess the cause of disease or health statuses. But, as already pointed out in Unit 1, because of the complexity of the relationships between risk factors in the real world, epidemiologist can never "prove" that a causal relationship exists between two epidemiological variables ("X causes Y").

In most cases, what epidemiologist can do is to observe that two epidemiological variables are associated (i.e. knowing the value of a variable in an individual "says something" about the value of the other) and, excluding spurious explanations for the association, conclude that the reason of the association is that the exposure causes the outcome.

For example suppose we observe that in a sample of individuals the prevalence of lung cancer is 5 times higher among smokers than among non-smokers ("smoking and lung cancer are **associated**"). Does it mean that smoking causes cancer? Not necessarily! But if we can exclude that the association is due to extraneous reasons (e.g. bad luck in choosing the sample, mistakes in measurements), we can "support" this causal explanation as the only plausible one.

Learning outcomes

* Understand association vs. causation
* Apply some basic criteria to assess probability of causation

Timing

You should be able to work your way through this session and the prescribed reading and task within 1 hour.

## Understanding association and causation

When epidemiologists observe that an association exists between two variables (an exposure X and an outcome Y), before concluding that the reason of this association is that X causes Y they must exclude alternative explanations.

Besides being the results of the exposure causing the outcome, an observed association between X and Y may in fact be due to the effects of one or more of the following:

* Chance (random error)
* Bias (systematic error)
* Confounding

**Therefore, an observed statistical association between a risk factor and a disease does not necessarily lead us to infer a causal relationship.** *Conversely,* **the absence of an association does not necessarily imply the absence of a causal relationship**.

The judgement as to whether an observed statistical association represents a cause-effect relationship between exposure and outcome requires that the researchers excludes (with reasonably certainty) each of the alternative explanations.

Statistics is the main tool used to exclude that the association is due to chance (i.e. "bad luck" in selecting a sample which is not representative of the population). In the previous Units we learned how to "quantify" the role of chance, calculating p-values and Confidence Intervals: if our p-value is small (or CIs are narrow and do not include the null value) we can "trust" our results and exclude that they are only due by chance.

But this is not enough. The reading by Gordis explains how the role of bias and confounding can be reasonably ruled out, it summarises the concept of sufficient cause and presents a series of well-known guidelines for judging whether an association is causal. The guidelines presented in the reading are usually known as the *Bradford-Hill criteria for causality*[[1]](#footnote-1).

**Reading**

Gordis, L. (2009). Epidemiology(4th edition). Philadelphia: Sanders. p 230-239.

A further criterion is usually added to those presented in the reading: the "strength" of the study design. This criterion says that there is a hierarchy in study designs (often graphically summarised as in the figure below), and results from studies higher up in the hierarchy (toward the top of the pyramid) being perceived as offering greater evidence of causality.

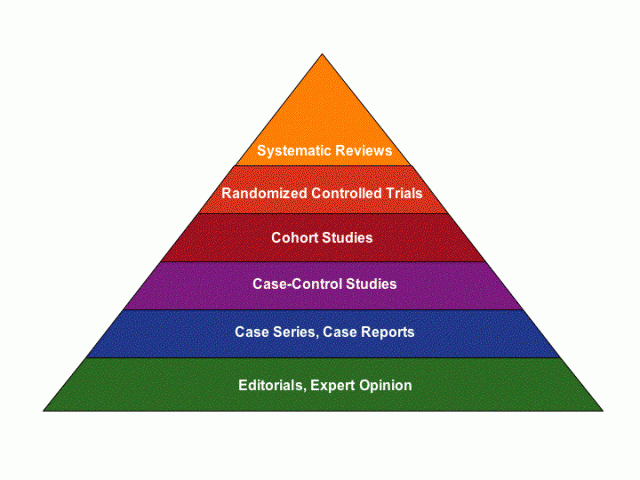


Figure: Hierarchy of study design

As a concluding remark, it is important to notice that the Bradford-Hill criteria (and others available in literature) are only guidelines, and, besides temporality, none of them is a necessary condition for the existence of a causal relationship.

No single criterion is completely reliable alone: the evidence provided by each criterion may be conflicting, and the likelihood of a casual relationships increases if different types of evidence point to same conclusion.

**Task 1 – Thinking about causality**

1. What is causal inference?
2. A statistically significant association has been demonstrated in a case-control study between the use of a drug for lowering cholesterol levels and risk of dying from heart attack. What more do you need to know before you recommend the withdrawal of the drug?
3. During an outbreak of severe neurological disease of unknown cause, it is suggested that the cause is adulterated cooking oil of a particular brand. Suppose you are an Epidemiologist and you are asked to decide if the culprit of the disease is indeed the cooking oil. How would you proceed? Which steps would you follow in your investigation? What type of study would be suitable? How would you assess the presence (or absence) of a causal relationship between cooking oil consumption and neurological disease from the results of your study?

**Task Feedback**

1. Causal inference is the process of determining with reasonable certainty that an observed association between two epidemiological variables is due to a cause-effect relationship.
2. First we need to exclude that the association, other than being statistically significant (i.e. not due to chance) is not spurious (i.e. not due to the confounding effect of extraneous variables). Then, If we have a reasonable certainty that the association is not spurious, we need to assess that the association is causal, i.e. that the drug causes death (for example, using the Bradford-Hill criteria).
3. There is no single way to conduct this type of investigation, and the objective of this task is to learn how to reason about causality (not to apply mechanical rules). Your tutor will comment personally on the results of your analysis

# Unit 5 – Session 2: Effect modification

## Introduction

When a presumed causal relationship exists between an exposure X and an outcome Y, it could be that the effect of X on Y is different according to the values of another variable Z. For example, the assumption of food containing gluten (exposure X) causes a series of negative symptoms, like diarrhoea and discomfort in the digestive tract (outcome Y) among individuals with coeliac disease but not among individuals without it. We say that the presence of coeliac disease (variable Z) modifies the effect of gluten digestion in the digestive tract. Coeliac disease is an **effect modifier** of the association between gluten and digestive tract problems.

This session shows how the presence of effect modification reflects on epidemiological data and how it can be assessed dividing the sample in sub-groups according to the values of the modifier, and analysing them separately.

## Learning outcomes

* Understand the concept of effect modification;
* Know the difference between multiplicative and additive interaction;
* Recognise effect modification looking at the results of stratified analysis.

Timing

This session should take you about 2 hours. A calculation based task is given- try to work your way through the calculations before looking at the feedback.

## The concept of effect modification

Effect modification is present when an exposure-disease relationship is different within different levels of another variable (the effect modifier). For example, a specific disease may be associated with an exposure in males but not with the same exposure in females.

Effect modification is something that one should be aware of and explain in the data. When effect modification is present, there is no meaning in calculating a global measure of association for the whole population, because the effect is different in different strata.

For example, it is known that women exposed to HPV (Human Papilloma Virus) have an increased risk of developing cervical cancer (there is strong evidence that HPV causes cervical cancer in women). The same relationship, obviously, does not exist for men[[2]](#footnote-2). In this case the variable "gender" acts as an **extreme effect modifier** (it cancels the effect of HPV in producing cancer in men), and the calculation of an average risk ratio of cervical cancer for a whole population including men and women is meaningless (and misleading). In these cases (and in less extreme examples) the correct procedure is to calculate and report **separately** the measures of association for men and women

The reading by Gordis presents in a more formal way the concept of effect modification (which is also known as **interaction[[3]](#footnote-3)**) and how this can be assessed looking at the study results stratified by different values of the (alleged) effect modifier.

**Reading**

Gordis, L. (2009). Epidemiology(4th edition). Philadelphia: Sanders: 256-261.

**Task 1 – Working with effect modifiers**

Of the 600 children between the ages of 4 and 18 living in an area of rural western Pennsylvania during an outbreak of Infectious Hepatitis (IH) in the early 1950’s, 66 developed IH during the period of the study.

The following tables gives the number of children contracting and not contracting the disease for two levels of socioeconomic status (low SES vs. high SES), and the same distribution stratified by age group (4-9 and 10-18).

a. Is there an association between SES and IH in the whole population?

b. Do you have evidence that age is action as an effect modifier of the association between SES

and IH?

Data from <http://practice.sph.umich.edu/micphp/epicentral/effect_measure_activity.php> [Accessed 13/11/2013]

|  |  |  |
| --- | --- | --- |
|  | Low SES | High SES |
| IH  Yes | 42 | 24 |
| IH No | 258 | 276 |
| Total | 300 | 300 |

**Table 1: Total sample**

|  |  |  |
| --- | --- | --- |
|  | Low SES | High SES |
| IH  Yes | 20 | 7 |
| IH No | 78 | 133 |
| Total | 98 | 140 |

**Table 2: Age group 4-9**

|  |  |  |
| --- | --- | --- |
|  | Low SES | High SES |
| IH  Yes | 22 | 17 |
| IH No | 180 | 143 |
| Total | 202 | 160 |

**Table 3: Age group 10-18**

**Task Feedback**

a) The risk ratio of IH for subjects of low SES compared to children with high SES is:

Therefore we can say that there is a (relatively weak) association between socioeconomic status and acquisition of IH: children with low SES have a higher risk of contracting IH.

b) If we calculate the RR separately by age group we obtain the following result:

|  |  |
| --- | --- |
|  |  |

This means that, among the youngest subjects, the risk of contracting IH is double for children with low SES compared to children with high SES. Therefore, in this age group, low SES is a relatively strong risk factor for IH.

However, in the second age group the risk ratio is 1: low SES does not affect the risk of contracting IH in this group. In other words, age is clearly[[4]](#footnote-4) acting as an effect modifier of the effect of low SES on acquisition of IH.

Unit 5 – Session 3: Generalisability of epidemiological inference

# Introduction

This last session of Unit 5 deals with the fundamental question which logically precedes any kind of interpretation of the results of epidemiological studies: are the results valid? Or, in other words, do the results represent the reality?

In this session we will define formally what we intend in epidemiology by "validity", and we will differentiate between internal and external validity.

## Learning outcomes

* Define validity of an epidemiological study;
* Differentiate between internal and external validity.

Timing

This is a short session and you should complete it within an hour.

1 Validity in epidemiological studies

When conducting an epidemiological study one would have to ***Choose*** one of these study designs and then appropriately ***Use*** the study design. Which study design one chooses depends on what research problem one wants to investigate within the constraints of available resources, available time, ethics and logistics. One should ***Choose*** the research design which is most able to ***Validly investigate*** the research problem and yet is **Feasible**. There is likely to be a trade-off between these two factors **(Validity and Feasibility)** when choosing, as the most valid study design might be the least feasible and then one must compromise and settle for a less valid study design. Or one must access more resources to make it more feasible, or one must change the research problem. However, sometimes one gets lucky and the most valid research design for the research problem is also the most feasible.

Epidemiological studies are conducted in a defined sample of subjects, usually randomly drawn from a larger population. Applying adequate epidemiological methods, the researchers draw some conclusions, usually in terms of the existence (or non- existence) of a causal relationship between two or more variables.

The reader of an epidemiological report might meaningfully ask in this regard two kinds of questions:

1. Do the results of the study represent the "reality" of the population which was studied?
2. Can the results be applied to a different population? If so, which populations?

The first question relates to what is called internal validity of a study (i.e. the extent to which the study results are "true" in the studied population), while the latter question relates to the external validity (or generalisability) of a study (i.e. the extent to which they can be generalised outside the actual population studied.)

Internal validity and external validity are two different but strictly interlinked concepts. In particular, internal validity is a pre-condition for external validity. The results of a study cannot be generalised to other populations if they are not even valid for the specific population which has been studied. The opposite is, however, possible. For example, a very well conducted clinical trial can provide internally valid results, but if the inclusion criteria for the selection of participants are extremely strict, the generalization to a broader population could be unwarranted.

## The reading from the Dictionary of Epidemiology, edited by Porta, gives a more formal definition of both internal and external validity.

**Reading**

Porta, M. (2008). A Dictionary of Epidemiology. New York: Oxford University Press: 252-253

1. Named after **Sir Austin Bradford Hill**  (1897 –1991), English epidemiologist and statistician who pioneered the randomized clinical trial and, together with Richard Doll, was the first to demonstrate the connection between cigarette smoking and lung cancer. His guidelines were published for the first time in the Proceeding of the Royal Medical Society (Proc R Soc Med. 1965 May; 58(5): 295–300). [↑](#footnote-ref-1)
2. Men can be infected by HPV, but they cannot have cervical cancer (!), therefore the risk ratio is 1 by definition: HPV does not affect the presence of cervical cancer. [↑](#footnote-ref-2)
3. More precisely interaction is the mathematical/statistical counterpart of the phenomenon of effect modification, but the terms are often used interchangeably in the epidemiological literature. [↑](#footnote-ref-3)
4. We are overlooking, for sake of clarity, the need to exclude that the result is due only to chance. In more advanced courses you will learn some statistical methods that can be applied at this end. [↑](#footnote-ref-4)