Cause and Effect in Epidemiology
Transcript

Cause and Effect in Epidemiology

Welcome to "Cause and Effect in Epidemiology." My name is Victoria Holt. As a nurse, I've worked in a variety of hospital and clinic practice settings, including public health clinics in East Tennessee and North Carolina. More recently, as an epidemiologist, I'm a faculty member at the Northwest Center for Public Health Practice at the School of Public Health and Community Medicine at the University of Washington in Seattle.

   For the last 15 years, I have also been a faculty member in the Department of Epidemiology at the University of Washington, where I currently teach courses in epidemiologic methods.

Cause and Effect in Epidemiology

Victoria Holt, RN, MPH, PhD

Northwest Center for Public Health Practice · www.nwcphp.org

About this Module

I'd like to mention a few points that may help make your learning experience more enjoyable.

   This module and others in the epidemiology series from the Northwest Center for Public Health Practice are intended for people working in the field of public health who are not epidemiologists but who would like to increase their familiarity with and understanding of the basic terms and concepts used in epidemiology.

   Before you go on with this module we recommend that you become familiar, if you haven't already, with the material presented in the following modules, which you can find on the Center's Web site:
- What is Epidemiology in Public Health?
- Data Interpretation for Public Health Professionals
- Study Types in Epidemiology
- Measuring Risk in Epidemiology

We introduce a number of new terms in this module. If you want to review their definitions at any time, the glossary in the attachments link at the top of the screen may be useful.

About this Module

Intended audience
People working in the field of public health who would like to increase their understanding of the basic terms and concepts used in epidemiology.

Recommended background
Familiarity with material presented in the following NWCPHP modules:
- What is Epidemiology in Public Health?
- Data Interpretation for Public Health Professionals
- Study Types in Epidemiology
- Measuring Risk in Epidemiology
   (See the Resources for links to these modules)

Our glossary of epidemiologic terms may be useful.

Course Objectives

This course offers an overview of the definition and aspects of the concept of cause and effect (or causality as epidemiologists would refer to it). By the end of this 45-minute module you should be able to define and describe the

Course Objectives

By the end of this module you should be able to
- Describe and distinguish between association and causality in epidemiology
- List and describe features of associations that support inferences of causality
- List principal concerns in inferring causality

infer: to draw a conclusion based on evidence

concepts of association and causality in epidemiology and distinguish between them. You should also be able to list several features of associations that support inferences of causality, and describe these features. Finally you should understand and be able to list several important or principal concerns that arise in inferring causality from epidemiologic studies.

Before we go on, I'd like to mention that this topic, causality in epidemiology, is often also called causal inference. To epidemiologists the term infer means to draw a conclusion based on evidence.

Importance of Causal Inference in Public Health

Why should you care about causality, or causal inference? Simply put, it's not just a topic of concern to epidemiologists. It forms the basis for making many important decisions in public health practice.

In a variety of situations or settings, public health professionals are called on to distinguish between association and causality, and this distinction—and subsequent actions taken as a result—may have far reaching implications for the public's health.

Importance of Causal Inference in Public Health

Why should you care?
Forms the basis for decision-making in a variety of public health practice settings
· Outbreak investigations
· Public health surveillance
· Disease clusters
· Public health program development

Adapted from Mesothelia in Montana, Montana DPHHS report, 2005 and Office of Vital Statistics and Montana Tumor Registry, Montana DPHHS

cause/agent ----▶ disease
cause/agent ───▶ disease

lung cancer
heart
low birthweight

Mesothelioma cases in Montana by county, 1979–2002

Mesothelioma cases
0   1-5   6-10   11-28

To name just a few examples: When outbreaks of infectious disease occur, there usually is an urgent need to identify the source or cause of the problem as a basis for developing and implementing control measures. In this situation it's important to distinguish between factors or agents that are merely correlated with disease and those that are truly causal, the removal of which is essential to halting the outbreak.

Understanding causes of disease may influence many public health surveillance activities beyond outbreak investigations. For example, if we know that smoking is a cause of lung cancer and heart disease and low birthweight, we might consider that information to decide to routinely monitor the prevalence of smoking in our community.

A disease cluster is defined as a greater-than-expected number of health events occurring within a group of people in a geographic area over a period of time. Clusters can involve either infections, diseases, or non-infectious diseases. We've already mentioned the usefulness of causal inference in investigating infectious disease outbreaks. And it is useful in non-infectious disease situations as well.

Although confirmation of a cluster of a non-infectious disease such as cancer does not necessarily mean that there is a single, external cause that can be addressed, knowledge of established causes of cancer and their prevalence in the community can be helpful in cluster investigations.

And finally, successful public health program development and imple-

2
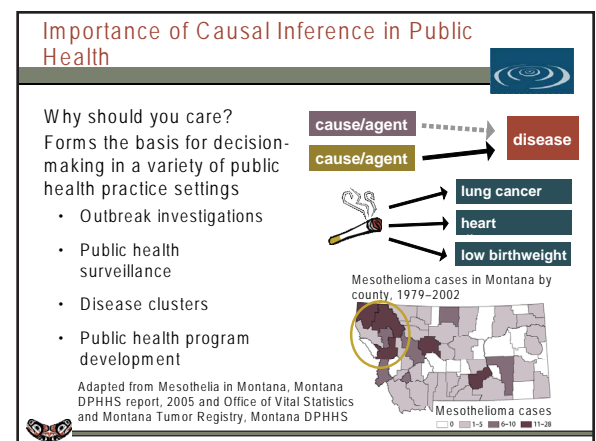
mentation rely on the identification of true causal factors that increase the risk of negative health outcomes in the community, in order to minimize the community's disease burden by targeting these factors.

Causal inference was the first step in a variety of notable epidemiologic accomplishments, such as decreasing coronary heart disease, mainly by decreasing smoking, high blood pressure, and cholesterol levels in the population.

Now let's turn to the topic of association in epidemiology.

Association in Epidemiology

Epidemiologists often talk about associations between variables. What we mean by association, in a general sense, is that there is a relationship or a connection between a certain exposure and a certain disease or health event. In other words, an association exists in a situation in which the frequency of the disease differs based on the presence or absence of the exposure of interest. Other names for exposure you'll see epidemiologists use are factor, risk factor, characteristic, or attribute.

A positive association means that in the presence of the exposure or risk factor we see a higher disease risk than we do in the absence of the exposure. This difference in disease risk is often measured by epidemiologists using measures of association called the relative risk and the odds ratio. If a positive association exists, the relative risk or the odds ratio will be greater than 1. A classic example of a positive association is smoking and lung cancer. Epidemiologic studies have shown that people who smoke are more likely than nonsmokers to be diagnosed with lung cancer.

A negative association occurs when the presence of the exposure or risk factor is seen with a lower disease risk. One example would be exercise, if we define regular exercise as the exposure under study. Many studies have found that people who exercise regularly are less likely than sedentary people to develop heart disease. In a negative association, the relative risk or the odds ratio will be less than 1.

For more information about the calculation and meaning of relative risk and odds ratio, see the module on Measuring Risk in Epidemiology.

Causal Association in Epidemiology

Epidemiologists use a definition of cause, or causal association, that's a bit different from that used historically in other disciplines. In epidemiology we say that a cause is a factor that plays a role in producing an occurrence of the disease. It just plays a role; it's not a necessary part of the disease

Association in Epidemiology

Associations between variables

Association: The frequency of disease differs depending on the presence of the exposure under study.

Positive association: The presence of the exposure is associated with higher disease risk.

- Relative risk or odds ratio > 1

  People who smoke are more likely than nonsmokers to be diagnosed with lung cancer.

Negative association: The presence of the exposure is associated with lower disease risk.

- Relative risk or odds ratio < 1

  Those who exercise regularly are less likely than sedentary people to develop heart disease.

process. For instance, we can talk about smoking being a cause of lung cancer even though some people who have never smoked also get lung cancer—smoking is not a necessary factor for all cases of lung cancer.

In the most general sense, a cause is something that if it weren't there, some cases of the disease wouldn't happen. This definition allows that factors can play a direct role or an indirect role in causing disease. A factor may not be capable of causing disease all by itself; it may be just one part of a more complex mechanism that necessarily involves other exposures or factors. For instance, not all smokers get lung cancer—smoking is not sufficient all by itself to cause lung cancer in all smokers. But we still consider smoking to be a cause of lung cancer.

The key feature of the notion of cause and causality is that causality is not proven in any one study. It's a process of determination or decision-making or inference based on a variety of information, as we'll discuss for the rest of this module.

> ### Causal Association in Epidemiology
>
> Epidemiologists use a definition of **cause** that is different from other disciplines.
> - Cause is a factor that **plays a role** in producing an occurrence of the disease.
>   - The causal factor is not a necessary part of the disease process.
> - Cause is something that if it weren't there, some cases of the disease wouldn't happen.
>   - The causal factor can play a direct or indirect role in causing disease.
> - Causality is not proven in any one study.

Causality Terms

Let's talk for a moment about some terms with specific meanings to epidemiologists.

Again, as a reminder, we observe associations—they are the results of specific studies. And we infer causes through a process of decision-making that often uses the guidelines we'll cover later in this module.

An observed positive association, such as between smoking and lung cancer, could lead us to an inference of causation. We would then say that smoking increases risk of lung cancer, that is, smoking is a risk factor for lung cancer.

An observed negative association, such as that between exercise and heart disease, could lead us to an inference of protection. We would then say that exercise decreases risk of heart disease, or protects against heart disease.

These statements, specifically the use of the words risk factor and protective factor, imply that you have made a decision about the causal nature of the relationships between the exposures and the outcomes under study.

Now we will pause for the first of several interactive exercises about the material we have just covered. Please note that the exercises sometimes take several seconds to load.

> ### Causality Terms
>
> Associations are **observed**.
>
> Observed positive association → Inference of causation
> - Smoking **increases risk** of lung cancer.
> - Smoking is a **risk factor** for lung cancer.
>
> Observed negative association → Inference of protection
> - Exercise **decreases risk** of heart disease.
> - Exercise **protects against** heart disease.

Exercise 1

Transcript

Causal Inference Guidelines

Now lets talk about guidelines epidemiologists use for causal inference.

First, and foremost—it's essential that an association be observed in order to proceed along the path of determining whether there's a cause-and-effect relationship. So let's say we observe an association between a certain exposure and a certain disease, and we want to know if that exposure truly is a cause of that disease.

Since causality is not proven in any one study, how do we determine if an exposure causes a disease? This is an important decision for public health practitioners to be able to make, as it may be the basis for determining whether to mount a campaign to decrease this exposure in a community.

This list of guidelines may help structure your thinking about the meaning of observed associations, to help you decide whether you can infer causality in specific situations. In the rest of the module we will discuss these seven guidelines:

1. Randomized trial evidence exists
2. No alternative explanations exist (or, as epidemiologists say, there is no confounding)
3. The timing of the relationship is correct (that is, the exposure comes before the disease)
4. The association is strong
5. The association is biologically plausible (that is, we know what the mechanism might be)
6. Higher doses of the exposure lead to progressively higher disease risk
7. And finally, the observed evidence of the association is consistent.

Let's consider the first of these features: randomized trial evidence.

Causal Inference Guidelines

Essential that an association is observed.
Causality is not proven by any one study.

1. Randomized trial evidence exists.
2. No alternative explanations exist.
3. Timing of the relationship is correct.
4. Association is strong.
5. Association is biologically plausible.
6. Higher exposures lead to higher risks.
7. Observed evidence is consistent.

exposure ? disease
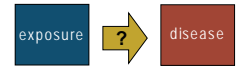
1. Randomized Trial Evidence Exists

The findings of randomized studies provide the strongest evidence pointing toward causality, because in these studies chance alone dictates which participants are exposed and which are unexposed.

In randomized trials a group of people is assigned to receive an exposure or an intervention, and these people are then followed over time to determine what proportion of them develop the target outcome under study, which could be an illness but could also be a beneficial outcome such as a decrease in blood pressure. At the same time, a group of people is assigned not to receive the exposure,

1. Randomized Trial Evidence Exists

Chance alone dictates which participants of the study are exposed.

Other factors don't distort the results.

Can't feasibly study all questions of causation with randomized trials.

Not ethical to use randomized studies for some types of risk factors.
  • Rely on observational studies.

Exposed

Outcome

No outcome

Not exposed

Outcome

No outcome

and this group is also followed up to determine how many develop the outcome, and whether that proportion differs from the proportion in the exposed group.

Because the researcher, rather than the participant, decides who will be in which group, other factors that could influence the risk of disease or the health event generally will not distort the results.

Unfortunately, we cannot feasibly study all questions of causation with randomized trials, and it isn't ethical to use a randomized trial to study some risk factors, such as suspected carcinogens. For these types of questions we must rely on observational studies such as case-control and cohort studies.

See the module on Study Types in Epidemiology for more information on types of observational study designs.

Randomized Trial Example

Here's an example of a randomized trial that has had wide-ranging public health effects. The purpose of the trial was to determine whether supplementation with folic acid or a mixture of other vitamins around the time of conception could prevent neural tube defects (which are serious birth defects that include spina bifida and anencephaly).

This study—called the medical research council vitamin study—was conducted in seven European countries in the late 1980s. All of the women enrolled in the trial had had a previous child with a neural tube defect and thus they were at high risk of another pregnancy complicated by this problem.

The women were divided into four groups and assigned to take different combinations of folic acid and other vitamins just before and during pregnancy. One group took folic acid supplements only, one group took folic acid plus other vitamins, one group took no supplements, and the final group took only the other vitamins.

The results of this study were striking. The nearly 600 women who were in the groups assigned to take folic acid (whether with or without other vitamins) had only six children with neural tube defects. The similar-sized group of women assigned to take other vitamins alone or no vitamins had 21 affected children. The relative risk was 3.5, and this difference was statistically significant. Thus, we would say there was a strong protective effect of folic acid supplementation on the basis of this study.

The authors concluded that "folic acid supplementation starting before pregnancy can now be firmly recommended for all women who have had an affected pregnancy, and public health measures should be taken to ensure that the diet of all women who may bear children contains an adequate amount of folic acid."
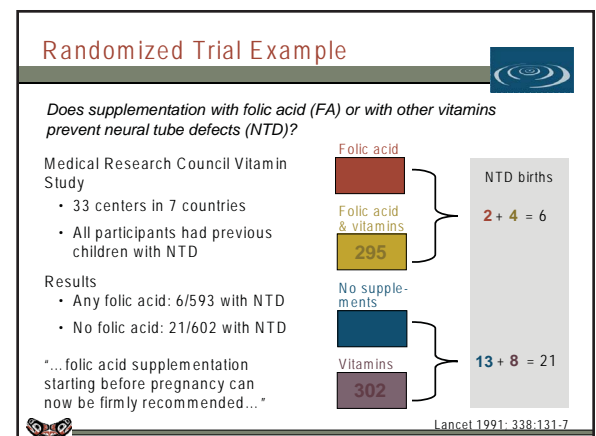
As a result of this and a few other similar studies, folic acid supplementation is now routine in pregnancy and some foods likely to be eaten by women of childbearing age are fortified with folic acid. You can see how persuasive a randomized trial can be in determining a cause and effect association.

2. No Alternative Explanations Exist

Our next guideline applies particularly in studies that are not randomized trials. We need to conclude that no alternative explanation exists for the association seen.

It's important to consider whether the increase in disease risk we see in the presence of a certain factor is due instead to other co-occurring factors or exposures. We call this situation confounding. Before we can infer a causal relationship, we first must consider the possibility of confounding and either dismiss it or take it into account. We can take confounding into account at either the design or analysis stages of a study.

What is confounding, exactly? It's a mixing of effects. If there is confounding, the association that we see between an exposure and a disease is a distortion. This distortion occurs because another factor or exposure that happens along with the one we're interested in is also associated with the disease we're studying—so what we're really seeing is a mixture of the effects of two exposures or factors on the disease.

In order for confounding to occur two things need to happen: First, there needs to be an association between the exposure you're interested in and this extraneous, or confounding, factor. And second, there needs to be an association between the extraneous factor and the disease you're studying.

Let's use an example to illustrate this principle.

2. No Alternative Explanations Exist

The increase in disease risk with a certain factor may not be due to the action of that factor (an alternative explanation could exist).

Confounding: A mixing of effects.
- Association seen between exposure and disease is a distortion

Exposure → Disease

Exposure → Factor
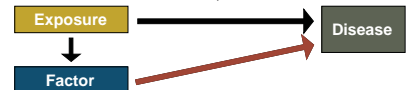
Factor → Disease

For confounding to occur there must be an association between:
- the exposure of interest and the extraneous factor
- the extraneous factor and the disease of interest

Confounding Example

Here we see the results of a study on whether alcohol drinking is a risk factor for lung cancer, that is, are people who drink alcohol more likely to get lung cancer than those who don't? This study also looked at whether, among those who drink, increasing amounts of alcohol consumed in a day leads to increasing cancer likelihood.

We see the crude, or unadjusted, results of the study here. It appears that alcohol intake is positively associated with lung cancer. We see that drinkers are more likely to get lung cancer, and those who drink the most heavily are the most likely to get it.

Confounding Example

Alcohol consumption and the risk of lung cancer

Alcohol intake (grams/day)	Lung cancer Incidence/10,000	Crude RR	Adjusted RR +
0	7.4	1.0	1.0
0.1–12	13.6	1.8	1.0
12.1–24	16.4	2.2	1.0
>24	25.2	3.4	1.1

Is there confounding by smoking?
- People who drink, more likely to smoke
- Smokers more likely to get lung cancer

+ adjusted for pack-years of smoking and other factors
See NWCPHP module on measures of association for the definition and calculation of a relative risk.

But before we decide that this is a causal association however, we have to ask whether confounding is responsible for these associations. You might specifically wonder if there is confounding by smoking. Why? Because smoking meets our first requirement for a confounder. An association exists between the exposure, alcohol drinking, and the extraneous factor, smoking. Research has shown that people who drink alcohol are more likely to smoke than people who don't drink. So we may have a mixing of the effects of alcohol and smoking when we think we're looking only at the effects of alcohol use.

And why might that matter? Because smoking also meets our other requirement for confounding—there's an association between smoking and the disease we're looking at, here, lung cancer. Smokers are more likely to get lung cancer than nonsmokers. So, the possibility exists that smoking is confounding the association between alcohol use and lung cancer. The effect we think is due to alcohol use may be due instead to smoking because many drinkers also smoke. To check this we conduct analyses that "adjust" for smoking, and when we do so, we see that there is no association between alcohol use and lung cancer. Drinkers are no more likely to get lung cancer than nondrinkers. Controlling or adjusting for the confounder, smoking, has removed the association between alcohol use and lung cancer, and therefore the association cannot be causal.

3. Timing of Relationship Is Correct

An essential feature of an association in order to be considered as causal is that the timing is correct. The suspected cause must come before the disease. Not only must the cause or exposure come before the effect or disease but there must be enough time for the suspected cause to have an effect.

The exposure-outcome time sequence can be difficult to establish, and it's easier in some type of studies than in others. In prospective cohort studies and randomized trials, exposures are noted at the beginning of the study, at a time when the study participants are determined to be free of disease, so it's easy to determine the time sequence. In contrast, in case-control or retrospective cohort studies records of exposures are obtained or subjects are interviewed about past exposures. In these studies it's not always clear whether the exposure of interest occurred before the disease process began.

It may also be difficult to determine the time sequence when the time period between the suspected cause and the effect is short. For example, Reye's Syndrome, which is a serious neurological disease, had an appar-

3. Timing of Relationship Is Correct

The suspected cause must come before the disease.
- Exposure-outcome time sequence can be difficult to establish.
  - In prospective cohort studies or randomized trials exposures noted at beginning of study.
  - In case-control or retrospective cohort studies past exposures considered.
- Can be difficult to determine time sequence when time period between suspected cause and effect is short.

Aspirin → 2-3 days → Reye's Syndrome
time
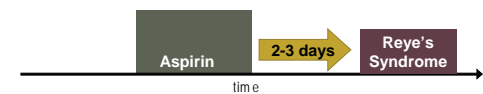
See NWCPHP module on study designs for more information on types of epidemiologic studies.

ent sharp increase in occurrence in children in 1980, and aspirin use was suspected as a cause. Researchers hypothesized that giving a child aspirin for flu or chicken pox increased the risk of the disease in the next few days. One concern raised early on about the proposed association was that children in the early stages of Reye's syndrome may have been given aspirin as a treatment for the disease, and thus one shouldn't conclude that any associations seen were cause-and-effect because the effect, Reye's Syndrome, may have come before the suspected cause, taking aspirin. Subsequent studies of this issue clarified that this wasn't the case.
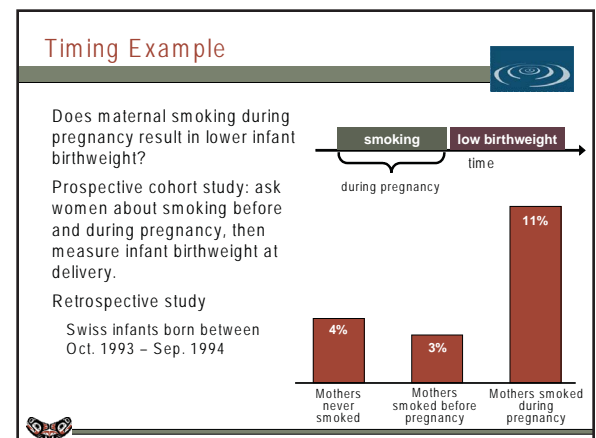
Timing Example

I'd like now to introduce an exposure-disease question that is of current public health importance. Does maternal smoking during pregnancy result in, or cause, lower birthweights? To answer this question we might first consider the idea of a randomized trial for a definite answer. However, in this case a randomized trial is unethical and impractical. Obviously we cannot tell some women to smoke during pregnancy!

So we turn instead to observational studies to address this question, and we must work a little harder to make causal inferences. The ideal observational study would be a prospective cohort study—in which women are asked about their smoking habits before and during pregnancy (perhaps several times) and then the infant is weighed at birth. This process ensures that the exposure (smoking) came before the effect (lower birthweight). Because of the availability of birth certificate data, studies of this question have often been retrospective in nature however—with information entered about smoking during pregnancy only after the birth has occurred.

One such study was done in Switzerland in the early 1990s. Using birth certificates, this study found that 76% of women who delivered during a one-year period had never smoked, and 4% of them had low birthweight infants. 5% of the women were ex-smokers who had smoked only before that pregnancy—and only 3% of them had low birthweight infants. Finally, 19% of the women were listed as smoking during pregnancy, and 11% of these women had low birthweight infants.

Because the effect of smoking on an infant's birthweight will occur before the infant is born, it's clear that the cause comes before the effect here, and this guideline for inferring causation is met. The fact that no effect on birthweight was seen among former smokers is further evidence that the important time period is during the pregnancy.

Let's pause now while you consider some questions about what you've just learned.

Timing Example

Does maternal smoking during pregnancy result in lower infant birthweight?

Prospective cohort study: ask women about smoking before and during pregnancy, then measure infant birthweight at delivery.

Retrospective study
Swiss infants born between Oct. 1993 – Sep. 1994

smoking → low birthweight
during pregnancy   time

- Mothers never smoked: 4%
- Mothers smoked before pregnancy: 3%
- Mothers smoked during pregnancy: 11%

Exercise 2

4. The Association Is Strong

Now let's return to our list of guidelines. Keep in mind that this guideline and all the following guidelines apply only to studies in which the exposure clearly comes before the disease and there is no confounding.

Perhaps the most intuitive guideline for causal inference is that of strength of the association. In evaluating the strength of the association between the suspected cause and the effect, the larger the value of the relative risk or the odds ratio (for a positive association), the less likely the association is to be false.

For example, one of the highest relative risks seen for a noninfectious disease was the scrotal cancer death risk of chimney sweeps. Even as late as the 1920s, chimney sweeps were 200 times more likely to die of scrotal cancer than were workers who were not exposed to tar or mineral oils. Remember, we always look for confounding factors, but it's almost impossible to imagine another explanation for such a large relative risk. If confounding played a part in the chimney sweeps' scrotal cancer death rate, the confounder would have to be extremely highly associated with both exposure and disease to account for the results seen.

Keep in mind two cautions about this guideline however. First, the strength of an association is not a biologically consistent feature—it depends on whether there are other causes of that disease and how common those other causes are.

And second, the relative risk or odds ratio does not necessarily have to be large to infer causality. Even an exposure with a modestly elevated relative risk could be a causal factor. For example, nutritional associations with disease are usually small, but consistent patterns emerge indicating causality is likely.

Strength of Association Example

What's the evidence that studies of maternal smoking and low birthweight meet this guideline of strength of association?

Using the Swiss study discussed previously, we can calculate the relative risk—it's the percentage of smokers with low birthweight infants divided by the percentage of nonsmokers with low birthweight infants—11 divided by 4—or 2.8. We would say then that smokers are nearly 3 times as likely as nonsmokers to have low birthweight babies. This may not be considered a strong association in a study, for example, of a food exposure and food-

4. The Association Is Strong

(This guideline and the following ones apply only when exposure comes before disease and there is no confounding.)

The larger the value of the relative risk, likely the association is to be false.

- Strength of association is not a biologic consistent feature; the size of the relati depends on the prevalence of other ca

- The relative risk does not have to be large to infer causality.

  - Nutritional associations with disease usually small, but patterns emerge indicating causality.

See NWCPHP module on measures of association for more information on the definition and calculation of a relative risk.
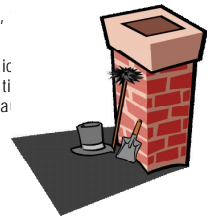
borne illness. Such studies often have relative risks that are much higher. For example, in an outbreak of salmonella infection in Oregon and British Columbia in 1995, a relative risk of 17 was found to be associated with eating alfalfa sprouts in the previous five days, in other words people who ate sprouts were 17 times as likely to become ill with salmonella as were people who hadn't eaten sprouts.

A relative risk of 2.8 isn't as strong as the risk in the salmonella outbreak, and it certainly isn't as strong as was the relative risk discussed just previously for chimney sweeps and scrotal cancer. But it's still a fairly strong association and would likely lead a reasonable person to keep considering the possibility that this relationship is causal.

Strength of Association Example

Association Between Smoking and Low Birthweight

Smokers are nearly 3 times as likely as nonsmokers to have low birthweight babies.

Relative risk: $\dfrac{11}{4}$ = 2.8

Relative risk = 17

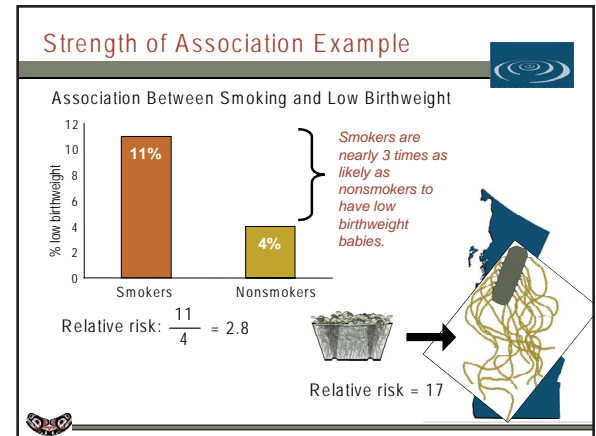## 5. The Association Is Biologically Plausible

Another guideline for causal inference is that the association you observe between an exposure and an outcome or disease should fit what we know about biology—or in other words, have biologic plausibility. If the association makes sense in terms of known biological mechanisms and processes, it becomes more likely that a cause-and-effect relationship truly exists.

A problem with this guideline is that biological plausibility can be difficult to judge—and what is considered plausible at any given time depends on what is known at that time. The association between Reye's syndrome and aspirin use, for example, was accepted as causal long before any biological explanation was evident, which shows that the lack of a known mechanism does not necessarily mean that the association is not causal. Unlike the guideline of the correct time sequence between exposure and outcome, this guideline is not an absolute.

5. The Association Is Biologically Plausible

If the association makes sense in terms of known biological processes or other epidemiologic knowledge, it becomes more likely as a cause-and-effect relationship.

- May be difficult to judge.
- What is considered biologically likely at any given time depends on the current state of knowledge.
- Lack of a known means does not necessarily mean that a particular relationship is not causal.

Aspirin → Reye's Syndrome

Biological Plausibility Example

What do we know about the biological plausibility of the relationship between smoking and low birthweight? Quite a lot, actually, and our knowledge enhances the likelihood that the relationship is causal.

Research has shown that smoking reduces blood flow to the placenta because nicotine causes vasoconstriction,

Biological Plausibility Example

What is the biological plausibility of smoking and low birthweight?

- Smoking reduces placenta blood flow (nicotine causes vasoconstriction)
- Carbon monoxide causes fetal hypoxia
- Cyanide causes fetal growth retardation

placenta
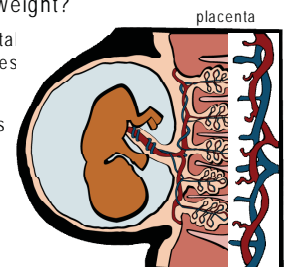
and that this reduced blood flow means less oxygen gets to the fetus, which interferes with growth.

Additionally, carbon monoxide, which is a by-product of smoking, also causes less oxygen to get to the fetus.

And finally, another by-product of smoking, cyanide, has been shown to slow fetal growth.

This body of biochemical evidence all points to the possibility that the relationship between smoking and lower birthweight is truly a causal one.

6. Higher Exposure Levels Carry Higher Disease Risks (Dose Gradient)

Another guideline that points us toward causality is that the association is strongest when we predict it would be. This is sometimes referred to as a dose-response gradient, or effect. Higher risks are seen with higher levels of exposure.

Although the observation of a gradient of disease risk associated with degree of exposure does point toward causality, keep in mind two cautions:

First, the presence of a dose-response relationship like this does not always mean that a cause-and-effect relationship actually exists. Confounding still could be responsible for the observed association.

And second, the absence of such a gradient doesn't mean that a causal relationship doesn't exist. We may not be able to see the dose-response relationship because of insufficient variability in levels of exposure. For example, most women in the US have a dietary fat intake of about 30 to 40% of total calories. There may not be enough women at very low levels of fat intake in the US to see the decrease in breast cancer rates that might occur with, say 15% of calories coming from fat.

Or a dose-response relationship may not be evident because there may be a threshold, an exposure level above which additional exposure has no additional effect.
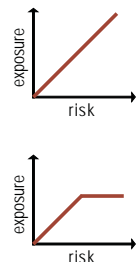
Dose Gradient Example: Lung Cancer

Remember this slide? We saw an apparent association between consuming alcohol and having an increased risk of lung cancer. Now notice that we also see an apparent dose-response relationship between alcohol use and lung cancer. Not only did drinkers have a higher apparent risk than nondrinkers, but the risk increased incrementally with the increasing amount of alcohol consumed in a

6. Higher Exposure Levels Carry Higher Disease Risks (Dose Gradient)

The association is strongest when predicted to be.

We can observe a gradient of risk associated with the degree of exposure.

- Presence of dose-response relationship does not mean that a cause-effect relationship exists.
- Absence of a dose-response gradient does not mean that a cause-effect relationship does not exist.

Dose Gradient Example: Lung Cancer

Alcohol consumption and the risk of lung cancer

Alcohol intake (grams/day)	Lung cancer Incidence/10,000	Crude RR	Adjusted RR +
0	7.4	(ref)	(ref)
0.1–12	13.6	1.8	1.0
12.1–24	16.4	2.2	1.0
>24	25.2	3.4	1.1

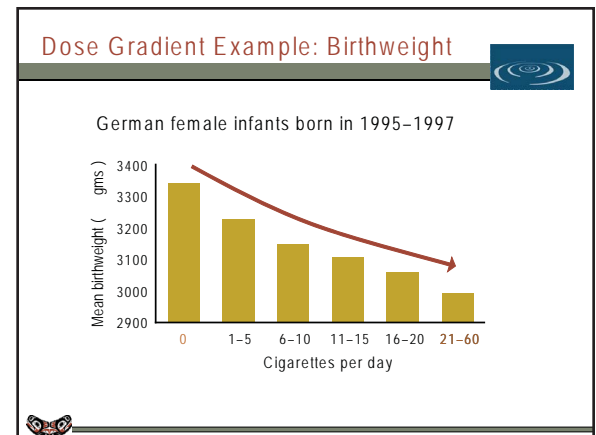+ adjusted for pack-years of smoking and other factors

day. But remember, the association disappeared when we controlled for the confounding effect of smoking. Relying solely on the existence of an apparent dose-response relationship to infer causality would be a mistake in this situation.

Dose Gradient Example: Birthweight

Here's an example of this dose-response guideline in relation to smoking and birthweight, from a retrospective cohort study done in Germany.

We see that among women who did not smoke during pregnancy, the average birthweight of their infants was more than 3300 grams, which is about 7 pounds 6 ounces. They are the unexposed group. And we see a fairly linear decrease in mean, or average, birthweight with increasing exposure doses, culminating in the lowest birthweight among women smoking 21 or more cigarettes per day. Their infants weighed less than 3000 grams on average, or around 6 pounds 9 ounces.

In this study the association was strongest when we predicted it would be so, among those with higher exposure to a toxic substance. This finding adds weight to the possibility that the association is truly causal.

Dose Gradient Example: Birthweight

German female infants born in 1995–1997

Mean birthweight (gms): 3400, 3300, 3200, 3100, 3000, 2900
Cigarettes per day: 0, 1–5, 6–10, 11–15, 16–20, 21–60

7. Observed Evidence Is Consistent

A final guideline for causal inference is the one that may be most used in practice.

Is there consistency of observed evidence? In other words, do repeated observations of this association exist under different study conditions?

If studies using different populations, different methodologies (case-control and cohort for example) and different circumstances, all have similar findings, this consistency points toward a causal relationship (in the absence, of course, of noncausal explanations such as confounding).

Sometimes, though, there are inconsistent findings. The lack of consistency can occur for various reasons having to do with the populations studied, the methods used, the differences in the way exposures are measured, and additionally, inadequate sample size can affect study results and their interpretation.

However, if studies in similar populations using similar methods and similar exposure measurements have diverse results, this inconsistency weakens a causal interpretation.

7. Observed Evidence Is Consistent

Repeated observations of the association exist under **different** study conditions.

- If studies with different populations, methods, and circumstances have similar findings, **consistency points toward a causal relationship**.

- Lack of consistent findings due to:
  - Populations studied
  - Methods
  - Exposure measurements
  - Sample size

- If similar studies have diverse results, **inconsistency weakens a causal interpretation**.

13

Consistency of Evidence Example

How consistent is the evidence linking maternal smoking to infant birthweight? Quite consistent. Evidence exists from studies of different types in different populations, and even, as we see here, using somewhat different ways to measure birthweight (mean, or average, birthweight and low birthweight percentage).

The first two studies seen here, done in Germany and Switzerland, were retrospective cohort studies that measured the association by looking at mean, or average, birthweights. They were remarkably similar. On average a 190- to 200-gram, which is 7-ounce, decrease in birthweight was seen among smokers' infants. The Swiss study also looked at low birthweight percentages, as did another retrospective cohort study done among low-socioeconomic populations in Liverpool England. The results of these two studies (relative risks of 2 and 3) were not identical, but both reported substantially increased low birthweight risk.

Finally, a different type of study (a case-control design) done in Washington, DC, among urban women who were primarily African American, found that women with low birthweight infants were twice as likely to smoke as those with normal birthweight infants.

So we see that in a variety of settings, using a variety of study designs, with two ways of measuring birthweight, the results are fairly consistent, and the message is the same: smoking in pregnancy adversely affects infant birthweight.

Let's pause now while you answer a few questions on what you have just learned.

Consistency of Evidence Example

Association of Smoking and Birthweight

Location	Study design	Findings
Germany	Retrospective cohort	Birthweights of smokers' infants were 200 gms lower
Switzerland	Retrospective cohort	Birthweights of smokers' infants were 190 gms lower RR for low birthweight = 3
England Low SES	Retrospective cohort	RR for low birthweight = 2
Wash. DC	Case-control	Women with low birthweight infants were twice as likely to smoke

Exercise 3

Summary

To recap a bit. We've talked in this module about the definition and aspects of the concept of causality as used by epidemiologists. An association is a relationship between an exposure (or risk factor) and a disease or health event. The frequency of disease differs based on the presence or absence of the factor.

A cause of a disease is something that plays a role in producing the disease. If the cause was not there, some amount of the disease would not occur.

Features of associations that support inferences of causality include: correct time sequence between exposure and disease

Summary

An **association** is a relationship between an exposure and a disease.

A **cause** of a disease is something that plays a role in producing the disease.

Features of associations that support inferences of causality include:
- ✓ Correct **time sequence** between exposure and disease onset
- ✓ **Strong** association
- ✓ **Biologically plausible** association
- ✓ **Dose-response gradient**
- ✓ **Consistent findings** in multiple observational studies
- Findings of association in **randomized trials**

Three principal concerns in inferring causality are the presence of:
- **Confounding**
- **Uncertain time sequence** between exposure and disease onset
- **Inconsistent findings** in studies with similar methods in similar populations

onset, strong association, biologically plausible association, dose-response gradient, consistent findings in multiple observational studies, and findings of association in randomized trials.

Three principal concerns in inferring causality are the presence of confounding, an uncertain time sequence between exposure and disease, and inconsistent findings in studies with similar methods in similar populations.

And we've seen these points illustrated using the example of the effect of maternal smoking during pregnancy on infant birthweight. The guide-lines of causal inference were met in investigations of this issue, namely that the association was seen in studies in which the exposure (smoking) clearly came before the outcome (lowered birthweight), the association was strong, with relative risks of 2 to 3, there are known biological mechanisms through which the exposure could plausibly operate, studies showed a dose-response gradient with more smoking leading to progressively lower birthweight, and finally observational multiple studies in different populations using different methods had similar findings. Obviously a randomized study wasn't possible. Thus, we have confidence that the association between smoking and low birthweight is causal, and any public health programs we develop with the goal of decreasing maternal smoking will have the potential to impact the frequency of low birthweight in our communities.

And this is but one example of the use of the process of establishing a cause and effect relationship as the essential first step in planning and imple-menting effective public health improvement programs.

Resources

If you would like to learn more about the concepts in this module, you might want to explore some of the resources listed here.

Now, if you're ready, please go on to the final assessment.

Final Assessment

Resources

Related online modules from the Northwest Center for Public Health Practice
- What is Epidemiology in Public Health?
- Data Interpretation for Public Health Professionals
- Study Types in Epidemiology
- Measuring Risk in Epidemiology

Other useful resources for further study
- Introduction to Epidemiologic Methods
  This multi-lecture series from the 2004 Epidemiology, Biostatistics and Clinical Research Methods Summer Session includes the following lectures:
  - Part 07: Causal Inference (one-hour video lecture)
  - Part 12: Confounding (one-hour video lecture)