Making Conclusions & Inferring Causality

EBCP Module #14
Outline of Module

This module will review or introduce the following topics:

- Bradford-Hill criteria
- Causal reasoning
  - Competing causes
  - Multifactorial causes
  - Effect modification
- Coherence between studies
Objectives of Module

- Students who complete this module should be able to correctly define and apply the following terms as they pertain to evidence-based clinical practice (EBCP):
  - Dose-response relationship
  - Coherence
  - Confounding
  - Multifactorial causes
  - Competing causes
  - Effect modification

- Students who complete this module should be able to:
  - Apply the Bradford-Hill criteria to evaluate causal explanations for study findings.
  - Explain the concept of effect modification.
  - Describe possible reasons for variations in evidence between studies.
Critical Appraisal

- In previous modules, you have been introduced to the idea of critical appraisal.

- We have suggested a method to order your appraisal of medical literature:

  1. Get your bearings
     - Identify question, design, etc.
  2. Evaluate internal validity
     - Bias (information and selection)
     - Confounding
     - Chance
  3. Evaluate external validity (generalizability)
Building on Critical Appraisal

- Our basic approach identifies the three primary threats to internal validity as:
  - Bias
  - Confounding
  - Chance

- However, there is more to making sense of research than just evaluating these common threats to validity.
  - This module will elaborate on some other considerations for your reading of the medical literature.
Causal Thinking as Clinical Thinking
Causal Thinking

- Causal thinking is fundamental to clinical practice.
  - It allows clinicians to know with varying degrees of certainty how illnesses present, what causes them, and how they can be cured.

- Causal reasoning should be supported by good evidence.
  - For example, well-conducted clinical trials (especially RCT’s) provide the best evidence regarding the efficacy of clinical treatments.

- However, it is not possible to use clinical trials to answer some important questions.
  - Cohort studies may provide the best evidence for questions where exposures cannot be assigned for logistic (e.g., gender or race) or ethical (e.g., a question about harm) reasons.
  - Case-control studies provide important evidence for questions where events cannot be repeated (e.g., determining the cause of a disease outbreak), where there are several possible exposures under study (e.g. genetic polymorphisms), or where the outcome is especially rare.
Causal Thinking

- As discussed previously, when possible, clinical trials (especially RCT’s) provide the best evidence for causation.

- Where trials are not possible, it is more difficult to argue for causality.
  - Several researchers and clinicians have suggested ways to help readers make sense of causality beyond the threats to validity that we have already described.
  - Here, we will consider the famous questions about causality proposed by Sir Arthur Bradford-Hill. They are sometimes called the *Bradford-Hill criteria*. 
Bradford-Hill Criteria
Historical notes on the Bradford-Hill Criteria

- To understand the Bradford-Hill criteria, it is useful to put them into context.
  - Sir Arthur Bradford-Hill was an important epidemiologist who, along with Sir Richard Doll, provided the evidence that smoking causes lung cancer, an association that was much debated during his time. He is also remembered as a pioneer of the clinical trial.
  - In 1965, he gave a speech to a group of physicians and researchers in the new specialty of occupational medicine.
    - The Bradford-Hill criteria come from this speech.
  - In the speech, Sir Bradford-Hill asks: “How do we determine what are physical, chemical and psychological hazards of occupation, and in particular those that are rare and not easily recognized?”
    - Although directed at questions about occupational hazards, his suggestions have been generalized to other areas of medicine and research.
  - The Bradford-Hill criteria are Sir Bradford-Hill’s suggestions about which aspects of an association should be considered when considering the case for causation.
Bradford-Hill Criteria

- Bradford-Hill suggested nine questions to help evaluate if an association is causal:
  - Strength: How big is the association?
  - Consistency: Is it seen in other studies (and in other times and places)?
  - Specificity: Is the association of exposure and outcome specific? Or does the exposure seem to be related to multiple outcomes?
  - Temporality: Does the exposure precede the outcome? Always? (This is critical to any assessment of causation.)
  - Dose response (also “biological gradient”): Is more exposure associated with more outcome?
  - Plausibility: Is the association biologically plausible?
  - Coherence: Is there evidence from the basic sciences to support it?
  - Experiment: Can the relationship be demonstrated experimentally through a trial in humans? (Experimental evidence is the strongest support for causation.)
  - Analogy: Are there similar associations in the literature for comparison?

- It is not important to memorize these criteria word-for-word, although you should be familiar with the ideas that they convey.
  - These questions will be useful in your critical appraisal of scientific literature, and you will run across them as you read medical research articles.
Bradford-Hill Criteria: Example

Bradford-Hill criteria are used frequently in clinical studies to provide support for causal explanations, particularly in the absence of clinical trial evidence.

An example is provided below.

Prior to the availability of clinical trial evidence on hormone replacement therapy and cardiovascular disease, researchers conducted a retrospective cohort study of women who had been followed for 13 years in a large health system.

Women were divided into groups of those who used non-birth control supplemental estrogen and those who did not.

The primary outcome of the study was mortality from cardiovascular cause.

Data from death certificates were used to ascertain cause of death. Secondary outcomes of the study were: all-cause mortality, cancer mortality, respiratory system-related mortality, digestive system-related mortality, accidents, homicide, and suicide.

The study was sufficiently powered to find meaningful differences.

Several multivariable analyses were conducted to account for confounders, including age, smoking, BMI, hypertension and socioeconomic status.

The authors of the study found lower cardiovascular mortality among supplemental estrogen users than among non-users.

- However, they also found that mortality in several other categories of cause of death was lower in users than in non-users—including for accidents, homicides, and suicides.
  - The authors state that the lower mortality in these three categories has no reasonable biological explanation.
- Although the associations that they found were statistically significant, the authors suggest that unaccounted for differences between users and non-users in their observational study may explain the finding about lower cardiovascular mortality.
  - They suggest that their study findings may have resulted from unaccounted confounding.
- They conclude that there is reason to doubt whether supplemental estrogen use causes decreased cardiovascular mortality and that this relationship in observational studies may be confounded.
  - They do not “prove” that supplemental estrogen use does or does not decrease cardiovascular mortality.
    - That hormone replacement therapy does not, in fact, decrease cardiovascular mortality (but actually increases it) was demonstrated in large clinical trials, including the Women’s Health Initiative (WHI), several years after the publication of this study.

To make their arguments about causality, the authors of our example study use several Bradford-Hill criteria, including specificity and plausibility.

- Similarly, by appealing to experimental evidence, another of the Bradford-Hill criteria, the WHI provided the best available evidence for the case that supplemental estrogen use does not decrease cardiovascular mortality.
Thinking about Causation

- As illustrated by the Bradford-Hill criteria, thinking about causation is not straightforward.
  - This is especially true when using evidence from observational (rather than experimental) studies.

- However, even in situations where experimental trials are well-conducted, the interpretation of evidence can be difficult.
  - Some other important aspects for you to consider when making conclusions about a scientific paper are laid out in the following slides.
Multifactorial Causes
Causal Thinking: Example

- Imagine a 72 year old woman who presents to your emergency room with a broken left hip. She reports that she was walking to her car that evening after a show at a theater. The streetlight was out and she did not see a patch of ice on the sidewalk. She slipped and landed on her left side.
  - What caused this woman to break her hip?

- Imagine:
  - If the streetlight had been on.
  - If the woman were some decades younger.
  - If the city had salted or sanded the sidewalk that night.
Causal Thinking: Example

- What caused the broken hip?
  - It seems reasonable that if any of the circumstances had changed—the streetlight, the ice, her age—the woman would not have broken her hip.
  - It is difficult to identify a singular cause without the others.

- We cannot always assume that if a single exposure occurs, a single outcome will follow.
Multifactorial Causes

- As in the previous example of the broken hip, causes can be multifactorial.
  - That is, they can require the presence of multiple factors.

- This is so common to clinical thinking that it is often taken for granted. For example:
  - Rh+ fetal blood alone does not cause maternal sensitization.
    - But, together with mom’s blood being Rh-, it can.
  - C. difficile spores alone are not sufficient for disease.
    - But, together with an immune-compromised state or long-term antibiotic use, they can.
  - Hepatitis D alone does not cause disease.
    - But, together with hepatitis B infection, it can.

- In all of the above cases, we need at least both of these factors together to cause the outcome.
Competing Causes
Competing Causes

Another important consideration is that multiple causes may effect the same outcome.

Consider, for example:
- What causes lung cancer? (Smoking, radon exposure)
- What causes fever? (Infection, cancer, crush injury, autoimmune disease)
- What causes stroke? (Hemorrhage, thromboembolism)

In these examples, unlike those for multifactorial causes, the exposures “compete” to cause disease.
- That is, each of the exposures listed can cause the same outcome all by itself.
Effect Modification
Effect Modification

- As we have shown, the relationship between cause and effect can be complex.
  - There may be multiple factors required in order to cause a particular outcome (multifactorial causation)
  - There may be multiple possible causes for a particular outcome (competing causation)

- As a final example of causal reasoning in medicine, we will discuss the concept of effect modification.
  - In statistics, this is sometimes known as “interaction.”
Effect Modification

- Effect modification occurs when researchers observe that the effect of an exposure differs with the presence of another variable.

- Effect modification is common in clinical practice. For example:
  - Age (effect modifier) modifies the effect of surgery (exposure) on recovery time (outcome).
    - Older people take longer to recover because they do not heal as quickly.

- You can see that this is not the same as previous examples:
  - The effect modifier is not simply part of a multifactorial cause. It is not required for the outcome to occur.
  - The effect modifier is not simply a competing cause, either.

- Remember, an effect modifier modifies the effect of the exposure on the outcome.
Effect Modification: Example

To better understand effect modification, consider the 2x2 table below for a hypothetical case-control study investigating the effect of alpha-1 anti-trypsin (A1AT) deficiency (a genetic trait) on emphysema among patients seen in a pulmonology clinic.

<table>
<thead>
<tr>
<th>A1AT Deficient</th>
<th>Emphysema</th>
<th>No Emphysema</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

200 100

We can calculate the crude* OR for the relationship of A1AT deficiency and emphysema.

Crude OR = [150/50]/[50/50] = 3.0

*used in this context, “crude” refers to the baseline OR that hasn’t been adjusted for any other variables.
In order to understand how smoking affects the relationship of A1AT deficiency and emphysema, the authors *stratify* the study subjects into two groups by their smoking status.

We can calculate adjusted ORs for the relationship of A1AT deficiency and emphysema controlled for smoking status.

- Among smokers, adjusted OR = \( \frac{120/20}{30/20} = 4.0 \)
- Among non-smokers, adjusted OR = \( \frac{30/30}{20/30} = 1.5 \)
The stratified analysis shows that smoking affects the relationship of A1AT deficiency and emphysema. In other words, **the effect of A1AT deficiency on emphysema is modified by smoking.** That is, the effect depends on smoking. It is greater among smokers and less among non-smokers.

In this example, smoking is an effect modifier. Effect modification occurs when the effect of an exposure differs with the presence of another variable. Effect modifiers are often very important factors clinically.

You may remember that we also used a stratified analysis to illustrate **confounding**. So how do we tell the difference between confounding and effect modification?

In the case of a variable that acts only as a confounder, the adjusted measure of association (OR, RR, etc.) would also be different than the crude measure.

However, with pure confounding, the adjusted measures in the stratified 2x2 tables are equal (or very close to equal).

Where there is effect modification, the adjusted measures in the stratified 2x2 tables are different.

This simple distinction is important to remember.
Effect Modification versus Confounding

- A variable may be only a confounder, only an effect modifier, or a variable may act both as a confounding variable and as an effect modifier at the same time.
  - The last option (both confounder & effect modifier) is the best description of how smoking acts on the relationship of A1AT deficiency and emphysema.

- We know that smoking acts as an effect modifier because the odds ratios, when stratified by smoking status, are different from one another.
  - That is, the effect of A1AT deficiency on emphysema depends on smoking.

- However, we also know that smoking causes the outcome (emphysema) and appears to be non-randomly associated with the exposure (A1AT deficiency) in this study.
  - This is consistent with the definition of a confounder (an outside factor associated in some way with both the exposure and the outcome)
  - To confirm that smoking is associated with the exposure, consider this 2x2 table that can be constructed from the stratified tables on the previous slide.

<table>
<thead>
<tr>
<th></th>
<th>Smoker</th>
<th>Non-smoker</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1AT Deficient</td>
<td>120+20=140</td>
<td>30+30=60</td>
</tr>
<tr>
<td>Not A1AT Deficient</td>
<td>30+20=50</td>
<td>20+30=50</td>
</tr>
</tbody>
</table>

OR=2.3
Therefore, smoking is related to the exposure. *We don’t have to understand why* in order for it to meet the criteria for confounding!
Blurring Category Lines

- It is worth noting that, sometimes, our view of a cause and effect relationship does not fit easily into a single category such as:
  - Multifactorial cause
  - Competing cause
  - Effect modifier

- For example:
  - Smoking (effect modifier) modifies the effect of asbestos (exposure) on risk of lung cancer (outcome).
    - Through its inflammatory effects on the lung, smoking leads to physiological changes that promote the harmful effects of inhaled asbestos fibers—it directly modifies the effect of asbestos.
    - However, smoking can also cause lung cancer without asbestos—therefore, it is also a competing cause.
      - As shown in the previous example, if a competing cause is also associated with the exposure group, then this variable may function as a confounder.

- The concepts presented here demonstrate that interpreting the results of clinical research may not always be easy or straightforward.
  - However, as you know, evidence-based practice is critical to the care of your patients.
  - Considering the Bradford-Hill criteria and other perspectives of cause and effect will improve your ability to critically appraise evidence used in the care of your patients.
Coherence in the Evidence
Clinical Realities and Multiple Studies

- You may find in clinical practice that your decisions are frequently not based on the results of single studies, but rather, on the “body of evidence” related to a question.
  - Studies usually do not stand alone in the scientific literature.
  - Generally, prior research has led to the studies that we evaluate and consider applying to patient care.
  - Further studies often follow these, as well.

- Scientific findings should be replicable.
  - As the evidence base expands, scientific ideas are proposed, tested, and re-tested.
  - This leads to multiple—and occasionally contradictory—findings about the same question.
Bodies of Evidence

- In future modules, we will consider specific methods used to summarize information from multiple studies, including:
  - Systematic reviews
  - Meta-analyses
  - Clinical guidelines

- At this point, you should be familiar with a few reasons that you might find differences between the results of studies of the same question. These include:
  - Differences affecting internal validity
    - Such as exposure and outcome measurement, differences in follow-up, and management of confounding
  - Multifactorial causation, competing causes, and effect modification
  - Random chance
    - Alpha of 0.05 allows a 1/20 chance of type I error
    - Beta of 0.20 allows a 1/5 chance of type II error

- Importantly, the fact that a finding consistently appears in the literature (one of Bradford-Hill’s criteria) is not a sufficient argument for causation.
  - In the example about hormone replacement therapy given earlier, dozens of observational studies had suggested that supplemental estrogen decreased cardiovascular mortality. However, these studies likely all had similar problems, such as unaccounted for confounding, that affected their internal validity.

- In future modules next fall, you will learn more about how to interpret bodies of evidence by building on what you have learned about interpreting individual studies.
Points and Distributions

- A final note on generalizing results from studies: As we attempt to generalize study findings to our patients, we must remember the differences between any particular patient and the sample that was used for the study.

- Points are not the same as distributions.
  - Is a particular patient average? Could the patient be an “outlier”?
  - *We certainly can’t guarantee that an individual patient will experience the “average” or even the “most common” outcome in a study!*
  - How do we decide what is “best” for a particular patient?

- These questions are not always easily or clearly answered. Yet, you will be faced with them frequently in clinical practice.
Key Points for Module 14

- Causal thinking is an important aspect of clinical reasoning. When you see an association, try not to assume causation unless there is good evidence for it!

- The Bradford-Hill criteria provide a list of questions that may help in evaluating the case for a causal relationship.

- Complex relationships between cause and effect include multifactorial causation, competing causation, and effect modification.

- *Effect modifiers* are third factors, other than an exposure or outcome, which modify the effect of the exposure on the outcome.

- It is possible for researchers to come to different conclusions when studying the same question – this may be due to factors related to internal and/or external validity.
Please complete the Module 14 quiz

Module Design Team: Martha Carvour, Matthew Rysavy, Timothy Bahr
Reviewed & Updated by T. Hegmann