Advanced Legal Epidemiology Methods

Full Script

Slide 1

Welcome to the Public Health Law Academy's training called **Advanced Legal Epidemiology Methods.** This training is brought to you by ChangeLab Solutions and the Centers for Disease Control and Prevention's Public Health Law Program. Special thanks to the Policy Surveillance Program of the Center for Public Health Law Research at Temple University Beasley School of Law for its contributions to the content development for this training.

Slide 2

The content developers of this training want to remind you that the information provided in this training is for informational purposes only and does not constitute legal advice. ChangeLab Solutions does not enter into attorney-client relationships. ChangeLab Solutions is a nonprofit organization that educates and informs the public through objective, nonpartisan analysis, study, and research. The primary purpose of this discussion is to address legal and policy options to improve public health. There is no intent to reflect a view on specific legislation.

Slide 3

Further, while every effort has been made to verify the accuracy of these materials, legal authorities and requirements may vary from jurisdiction to jurisdiction. Always seek the advice of an attorney or other qualified professional with any questions you may have regarding a legal matter.

The contents of this presentation have not been formally disseminated by the Centers for Disease Control and Prevention and should not be construed to represent any agency determination or policy.

Slide 4

Individuals who work as public health practitioners, lawyers, and policy experts in the field of legal epidemiology need measurable skills to move their careers forward. CDC's Public Health Law Program developed the Legal Epidemiology Competency Model in 2018 to help reflect the practice of legal epidemiology and guide practitioners' career trajectories.

This module of the Public Health Law Academy will cover the four tasks listed on this slide to build skills for entry-level (tier 1) professionals in the field of legal epidemiology. These four tasks are not the objectives for this course; they are general legal epidemiology competencies. The four tasks are

- Identifying opportunities for a legal epidemiology study to address legal, health, or other issues;
- Determining which legal epidemiology study designs address potential associations between law and health;
- Collecting and analyzing qualitative and quantitative study data using generally accepted research methodologies; and
- Interpreting results, drawing conclusions, and formulating key findings to help improve public health.

This training is intended for individuals who are not in management or supervisory positions and who directly implement legal epidemiology research tasks such as collecting, analyzing, and coding data.

Slide 5

This course will introduce concepts that are more advanced than those covered in other Public Health Law Academy trainings. It will focus on concepts of legal epidemiology studies, also known as *legal epi studies*. It is intended for

- People with a background or a graduate degree in public health, psychology, economics, or a related field;
- Researchers and practitioners with experience in quantitative or qualitative analysis; and
- Other experts who have used statistical or evaluation methods in their work.

Many of the concepts we'll discuss are complicated and may take time to absorb. Take your time in working through this material and thinking through how it might apply to your work.

Slide 6

In the first part of this training, we will provide a quick recap of legal epidemiology.

Slide 7

Next, we will introduce the three phases of legal epi studies.

Slide 8

In Phase 1, we will describe the steps in planning an advanced legal epi study. We will discuss logic models, three basic socio-legal theories, and where to find data for legal epi studies. We will also explore options for research designs.

Slide 9

In Phase 2, we will explore the advantages and limitations of various research designs.

Slide 10

And in Phase 3, we will talk about examining and understanding data.

Slide 11

But before we discuss these new topics, let's take a moment to review what we already know about legal epidemiology.

Slide 12

Legal epidemiology is the scientific study and deployment of law as a factor in the cause, distribution, and prevention of disease and injury in a population. It uses transparent, scientific methods and rigorous quality control.

In this training, we'll focus on advanced legal epi studies. Specifically, we will look at how scientific methods can help us measure processes and outcomes. As an example of how legal epi evaluation studies work in real life, we'll explore how to use advanced research methods to determine whether distracted driving laws affect health outcomes.

Slide 13

Let's talk about Wendy. Wendy is a public health expert who works for her state's public health department. Her task is to determine whether laws on distracted driving affect health outcomes. She has chosen to use policy surveillance to answer this question.

Slide 14

Wendy and her team conducted a legal mapping project. At the end of that process, Wendy's team had robust legal data. But that was just the first step. To fully understand the effects of distracted driving laws on health, Wendy must compare the legal data her team collected with public health outcome data.

Slide 15

Because the goal of Wendy's study is to find statistical relationships between the legal landscape and health outcomes, she decides to give this task to the researchers in her office who have experience in conducting quantitative analyses.

These researchers will analyze the two datasets to determine whether distracted driving laws have a positive relationship, a negative relationship, or no relationship with health data on motor vehicle injuries.

Ultimately, Wendy and her fellow experts are trying to answer the question "Do distracted driving laws work?" In this training, we will outline the steps Wendy needs to take to figure it out, and we'll show how you and your team can conduct a similar study.

You might not have the training to conduct this kind of analysis, using the data from your legal mapping project. If that's the case, ask someone with the right experience and skills to help you.

Slide 16

To understand the effects that distracted driving laws have on health, Wendy's team needs to answer a few questions:

- Which cell phone–related behaviors do distracted driving laws actually regulate? For example, do they address calling, texting, and web surfing?
- What are the penalties associated with these laws?
- Are these laws enforced?
- Do the laws reduce motor vehicle crashes?
- Do the laws reduce injuries and deaths caused by motor vehicle crashes?

Slide 17

The team's initial research questions will be used in planning, operationalizing, and analyzing data, which are the three phases of an advanced legal epi study. In this training, we'll define each phase, describe the steps within each phase, and provide examples for each step.

Slide 18

Let's begin with Phase 1. The next set of slides describes the steps involved in planning an advanced legal epi study. To help us along, we will continue to use Wendy's work as an example.

Slide 19

It's important to note that each phase is iterative. When conducting a legal epi study, you'll likely make some progress and then discover that you need to revisit something. For instance, you might need to go back to the evidence base to refine the question or the logic model. We will walk you through each of the steps in the planning phase:

- Reviewing the evidence
- Drafting research questions
- Identifying relevant laws and legal mechanisms
- Developing a logic model
- Checking data sources
- Selecting a research design

Slide 20

The first thing Wendy should do in the planning phase is check the scientific literature base for published studies examining the impact of distracted driving laws. Remember, her question is very general: Do distracted driving laws work?

To answer this question, she'll need to define the specific impact she wants to assess. This outcome must be observable and measurable. For example, she might measure the number of crashes, number of injuries, or number of deaths to determine whether distracted driving laws work.

Slide 21

After Wendy's team members have reviewed the evidence base, they should have a sense of what research questions they still need to answer.

At this point, the team should identify a specific research question that can be tested empirically. For instance, Wendy's research question might be "Does the presence of distracted driving laws reduce crashes?" The null hypothesis associated with this research question is "Distracted driving laws do not reduce crashes," and the alternative hypothesis is "Distracted driving laws do reduce crashes." But Wendy's team might pick something more specific. For instance, they might investigate the following research question: "Do distracted driving laws with higher fines reduce the number of crashes significantly more than distracted driving laws with lower fines?"

Slide 22

As noted on the slide, the team can take several different approaches. What is important is that they pick a specific law or feature of law and pair it with an implementation or enforcement activity (a legal mechanism) that can be measured. This activity will help them figure out which laws or provisions are really making a difference.

Let's use Wendy's work to explain this further. Wendy must first decide what she is investigating. Does she want to measure the effect of the presence of a distracted driving law as a whole? Or is she more interested in how a particular feature of the law affects health? After she picks what she wants to measure, she must examine how the law or feature of law has been implemented or enforced. For example, to measure enforcement, she could use the rates of people cited for texting and driving or the number of citations written at a randomized selection of road intersections.

Slide 23

It is tempting to try to study law as a single event or occurrence. For example, Wendy might try to study a distracted driving law's direct effect on motor vehicle crash rates in a certain state. But there are many variables that affect how laws play out in the real world.

So, what do we do? Because laws and their effects are complex, we use **logic models** to articulate theories of change and to describe visually how a law is supposed to work. Logic models can help determine what data are needed to test the hypotheses articulated in the logic model.

We're now going to spend some time exploring how a logic model can be used to study the effects of laws on health. The next set of slides will walk you through the overarching, macro-level logic model of legal epi studies. This logic model was created by the Center for Public Health Law Research at Temple University Beasley School of Law. Notice that there are boxes and arrows in this diagram that show how laws are intended to affect behaviors and environments, which in turn affect population health.

Slide 24

Let's start with how laws are enacted. Lawmaking involves many activities: bill writing and editing; introduction of the issue to legislative committees and to the general public via public notice; policy refinement; lobbying; executive review; and, finally, codification.

Slide 25

Lawmaking, then, leads to laws on the books. A law could be a statute, a regulation, case law, an executive order, an administrative rule, and so on.

Most interventional legal epi studies begin with laws on the books. Stakeholders want to know what effect the existing laws will have on intended outcomes. In our example, Wendy reviews the actual text of the distracted driving law in her state and in neighboring states. She analyzes all the observable features of those particular laws.

Slide 26

Let's pause for a moment. Take a minute to think about this question: Which features of a law might Wendy want to study?

Slide 27

If you said who is being regulated or who the law applies to, that's correct! She might also be interested in studying other features, such as

- What the penalties are. Are they fines, jail time, a revoked license?
- Where distracted driving is prohibited. Do distracted driving laws apply only in school zones, on highways, or on local surface streets?

Wendy might also want to examine what the enforcement plan is – for example, whether the police must ticket drivers who are using their cell phone or whether they ticket drivers only when they've committed another traffic violation. What might be some of the unintended consequences of the enforcement plan, especially on communities with low income and communities of color?

Slide 28

Now, let's take a look at legal practices. Legal practices are all the activities that occur after a law is passed, such as appropriating funding, implementation, and enforcement.

Let's refer back to our distracted driving example. Knowing something about enforcement is critical in this example. Generally, it is the police who enforce distracted driving laws. If the public feels no threat of penalty for texting and driving, then we are not likely to observe an effect of the law.

Slide 29

Let's review what we've just discussed. In sum, the lawmaking, the laws on the books, and the legal practices are all the pieces of law and policy that can be examined in a legal epi study. Here are some examples:

Slide 30

We could design a study of how a law gets passed, which is a study of lawmaking.

Slide 31

We could design a study of the effects of a law on the books without looking at how it is implemented or enforced, although this approach would not be ideal.

Slide 32

We could design a study of how laws are implemented differently across jurisdictions.

Slide 33

Or we could design a study of how laws are enforced.

Slide 34

Public health laws are aimed at changing individual behaviors and the environments in which people live, work, and play. The end goal of public health laws is to improve population health. Examples of population health outcomes include the number of car crashes and the rate of injuries due to car crashes; the incidence of vaccine-preventable childhood illness; the number of drug overdose or accidental poisoning deaths; and estimates of chronic disease prevalence.

Slide 35

But how can we tell whether a law makes a difference? Public health laws are written and passed to protect the public's health. In defining the public's health, we often need to think in intermediate steps. These laws are intended to directly change behaviors and environments, which in turn affect population health.

Slide 36

Let's revisit our distracted driving example. In this case, the laws on the books are likely intended to deter people from texting or talking on the phone while driving because they fear a violation and a fine.

Slide 37

The behavior affected is use of a cell phone while driving.

Slide 38

In a legal epi study examining implementation, Wendy and her team might study the rates at which the police issue violations and fines. This research could help them determine whether the law on the books is sufficient or whether enforcement is a key piece that makes the law work.

Slide 39

In the next three slides, we will explain the arrows in the overarching logic model, which are labeled with letters in this graphic. The arrows represent the causal mechanisms between the boxes. In other words, the arrows are what actually happens on the ground between a law being passed and a health outcome. To begin, we will examine causal paths A and B.

Slide 40

Path A: the arrow between lawmaking and laws on the books explores the factors that influence what laws are enacted as well as how they're enacted.

Slide 41

Path B: the arrow between laws on the books and legal practices explores how existing laws are transformed into what agencies are actually doing on the ground. For example, an implementation study might examine the practices of agencies that are putting a law into action.

Slide 42

Now let's dive a little deeper. When creating a logic model, we might have lots of data about environments or population health. But one of the harder things to understand – and therefore to measure – is the relationship between legal practices and changes in behavior, which we will now discuss.

Slide 43

Legal epi studies typically start with laws on the books. The studies then investigate the impact those laws have on behaviors or environments, represented here by Path C. These behaviors and environments are **mediators** of the ultimate measure of health itself – a relationship that is represented by Path E.

Slide 44

A **mediator** is an "in-between variable" along the causal pathway.

Slide 45

A mediator is affected by its inputs . . .

Slide 46

And, in turn, affects the outcomes.

Slide 47

Sometimes, legal practices, rather than laws on the books, can be the starting point for a legal epi study. An example of a legal practice that might be the start of a study is the number of citations issued to enforce a seat belt law.

Slide 48

It's worth noting that many studies limit themselves to Paths B and C.

Slide 49

Some studies follow only Path C. In these cases, the focus is on how the law or the law's implementation has led to observable changes in behaviors and the environment. These behavioral and environmental changes are used as proxies for changes in health outcomes.

Slide 50

For example, a study of the impact of tobacco warning labels might measure outcomes such as the smoking rate or tobacco sales, rather than deaths from lung cancer.

Slide 51

Path D captures the important relationship between environment and behavior.

Slide 52

For example, a seat belt law might work by inducing changes in individual behavior through fear of a fine . . .

Slide 53

But the widespread adoption of seat belt use by individuals can change social norms. This kind of large-scale shift can make other people who are not worried about fines adopt the behavior.

Slide 54

The final step – and the hypothesized long-term outcome – is **population health**. Population health is broadly defined here, and any number of outcomes can appear in this box. For example, a population health outcome for Wendy's study might be the number of motor vehicle crashes per capita in a state or the number of deaths due to crashes.

Slide 55

In the next section, we'll highlight three of the most important socio-legal theories that help explain why and how law changes – or does not change – behaviors. The three theories we'll examine are deterrence theory, procedural justice theory, and the theory of planned behavior.

Slide 56

Deterrence theory examines the use of legal levers, such as high fines, to deter people from doing something like texting while driving.

Slide 57

Procedural justice theory is about perceived fairness of the law. If individuals perceive that the law is applied fairly and equally to everyone, then they are more likely to comply with it.

Slide 58

The **theory of planned behavior** is in play when social norms are involved in changing behaviors. Think about the seat belt law we mentioned a few minutes ago, or consider a helmet law. When these laws were first introduced, very few people complied. Over time, people began to realize that seat belts and helmets really saved lives. As a result, social norms started to change, and more people began complying with the law.

Slide 59

Now we will expand on each of these theories, starting with deterrence theory. Legal mechanisms, or legal effects, are the ways in which a law works. According to Merriam-Webster's Dictionary, *deterrence* is "the action of discouraging an action or event through instilling doubt or fear of the consequences." Deterrence works by having a law on the books with a particular sanction, such as a fine or a prison sentence.

Deterrence theory suggests that before deciding whether they will comply with the law, people will assess the strength of the penalty and the likelihood that they will be caught and punished. If the penalty is high enough *and* if the law is enforced so that the likelihood of getting caught is high, people will be more likely to comply with the law.

Slide 60

Consider Wendy's scenario. For distracted driving laws to succeed in reducing crashes, drivers must perceive the penalty for getting caught as costly and the likelihood that they will be caught as strong. These perceptions will make people stop texting and driving. In Philadelphia, for example, a distracted driving law is on the books. But currently there is little enforcement and minimal threat of penalty, so the rate of texting and driving remains high.

Slide 61

Let's move on to procedural justice theory. This theory suggests that individuals are affected by how they perceive they will be treated under a particular law and by the law in practice, which might include implementation and enforcement.

Different individuals or groups might perceive and experience the implementation or enforcement of a law differently. If an individual feels that the law is legitimate and that it is implemented and enforced fairly across the population, then that person is more likely to comply with the law. However, if the law itself – or its implementation and enforcement – is perceived as unfair or illegitimate, then the likelihood of compliance is reduced.

Slide 62

An example of how procedural justice theory might work in practice lies in the inequality of police interactions across racial groups. For instance, Terry stops in New York City often unfairly target people of color. A Terry stop is the brief detention of a driver or pedestrian on reasonable suspicion of involvement in criminal activity. The stop does not lead to arrest because the police cannot show probable cause that the person they have detained has violated the law. When the police disproportionately stop drivers of a particular race or gender under distracted driving laws, the public understandably considers these laws to be unfair and are less likely to support the laws.

Slide 63

Finally, let's dive into the theory of planned behavior. This socio-legal theory suggests that people's attitudes about a law affect their intent to comply. Attitudes can be individual or collective. If subjective norms related to a law are negative, then the likelihood of the public complying with the law might be low. As subjective norms and attitudes toward the law begin to change, the likelihood of compliance changes as well.

Slide 64

For example, some laws authorize public awareness campaigns aimed at publicizing the dangers of texting and driving. Billboards showing crashes are intended to change people's attitudes about texting and driving so that when they get into a car, their intentions to text and drive are reduced.

Slide 65

As the graph on the slide indicates, compliance with older laws – such as no-smoking laws – is also affected by social norms. In the 1970s, the media depicted smoking as cool because the dangers were not yet fully known. As experts began to understand the hazards of smoking, the media started to document its health consequences among generations of families. The resulting shift in social norms on smoking contributed to acceptance of and compliance with more recent no-smoking and clean indoor air laws.

Using these socio-legal theories and historical examples of how they work as background, we will now use logic models to examine some current problems that have complex linkages between the law and changing behaviors and environments.

Slide 66

Let's proceed as if this is the first logic model you've created – and it might well be! It helps to start your thinking from the right-hand side of the model. First, identify a public health problem. In this case, let's pick the number of school days missed by all students in a school. Next, work your way back in the causal chain to identify the proximate causes and potential mediators.

Although it may not be immediately obvious, this example is rooted in the social determinants of health. We know that poverty and related conditions affect student attendance, but it is very difficult to jump directly from poverty to school absences. The challenge is to identify the proper causal chain. This is an iterative process driven by the evidence base, scientific theory, and input from stakeholders and experts.

Slide 67

Let's look at poor housing quality, for example. Poor housing conditions can increase a child's risk for asthma and in turn affect whether she feels well enough to go to school. Laws and legal practices can influence whether homes are built and maintained in ways that prevent or exacerbate asthma.

Can you identify laws that may play a role in the number of missed school days because they're related to housing quality?

Slide 68

Some of the relevant laws include the Fair Housing Act, mandatory mold remediation programs, housing court decisions, and housing code enforcement.

Slide 69

We're now ready to move on to the next step in the planning phase: checking data sources.

Remember, there are different types of data. Cross-sectional data represent a single point in time. Longitudinal data encompass data over time.

Slide 70

When you're first creating a hypothesis for a legal epi study, it is necessary to identify the unit of analysis, which defines what you're trying to study.

For legal data, the unit of analysis can be laws or policies from federal, state, or local governments or from tribes, territories, courts, or even school districts. Comparing like sources, such as jurisdictions at the same level of government, provides a legal unit of analysis.

For outcome data, the unit of analysis can be outcomes from individuals, institutions, or systems. For example, Wendy captured state-level laws and wanted to compare their content to the number of traffic fatalities attributed to distracted driving – including the number of car crashes, traffic fatalities, or hit-and-runs.

Slide 71

Well-designed studies of public health laws in real-world settings help promote effective regulatory strategies. But public health laws are often passed at the state or local level, so it is impossible to randomly assign individuals to comparison groups, as experimental research design would dictate.

So, when a law is passed in at least one jurisdiction, how might we examine the impact of that law? A good comparison would be a state without the law. And when other jurisdictions start to pass the same kind of law, often the features of the law are not exactly the same across jurisdictions. This situation creates a natural laboratory for studying laws and the features of laws.

In **natural experiments,** scientists do not control when and where "treatments" are implemented. Changes in laws, regulations, or systems that affect population health provide opportunities for natural experiments. Policy surveillance of texting-and-driving laws from 2005 onward can provide data for natural experiments. In August 2005, 18 states and the District of Columbia had laws restricting texting while driving. Those states were Oregon, California, Arizona, Colorado, New Mexico, Texas, Oklahoma, Arkansas, Louisiana, Mississippi, Tennessee, Kentucky, Illinois, Georgia, Delaware, New Jersey, New York, and Maine. The rest of the states did not have any laws restricting texting and driving.

Slide 72

Let's pause here for a moment. We've covered *a lot* of information, and it may take some time to absorb. Are there any questions?

Slide 73

In the first half of this training, we talked extensively about logic models and socio-legal theories. Now let's finish discussing the planning phase and then introduce the final two phases of advanced legal epidemiology studies: operationalizing research designs and analyzing data.

Slide 74

Picking back up with the last step of the planning phase, let's take a look at research designs. A true experimental design allows for causal inference, so the key feature of an experimental design is random assignment to treatment and control groups.

Why is random assignment important for causal inference? When you randomly assign subjects to treatment and control groups, you can assume that any difference between the two groups at the end of the experiment is attributable to the treatment. The assumption is that the two groups are comparable on all other observed *and* unobserved features at the start of the experiment. This is a very big assumption, and it is critical to understanding why random assignment is required for experimental designs.

Let's say we're studying the effects of a new pharmaceutical – "Drug A" – in reducing blood pressure. To do so, we take a random sample of males who are 25 to 55 years old. We then randomly assign half to the treatment group and half to the control group. When we do this, we assume that any differences between the men at the start – such as their diet, their level of exercise, and their genetic propensity for heart disease – are randomly sprinkled throughout both groups. This assumption applies to variables you measure and variables you don't; that is why you can claim you ruled out all other possible explanations.

In the lab, it is possible to randomly assign mice to a particular dosage of a drug. In clinical trials, it is possible to randomly assign subjects to varying treatment conditions and a control condition. But in the real world, it is rarely possible to randomly assign people to exposure to a law or policy.

Slide 75

A **quasi-experimental design**, which is also called a *natural experiment,* allows researchers to use available, observable data and correct for the absence of random assignment through the design. In a study of the impact of seat belt laws, for example, researchers have a number of tools they can use to properly test for causal effects.

One technique is the use of an appropriate comparison group instead of the treatment–control scenario. Ideally, a study would compare people who had been "treated" by the law with those who had not. To make this comparison, you would identify at least one state with the law and at least one state without it. Then, you would explore whether the laws or legal features are comparable and applied to similar populations. For example, a seat belt law can be assessed over time to find out whether more people wear seat belts and whether more traffic fatalities are prevented because of it.

Slide 76

The quasi-experimental design toolbox contains many tools:

- Repeated measures
- Time resolution
- Functional form of effects
- Comparison jurisdictions
- Comparison groups
- Comparison outcomes
- Replications
- Dose response

We'll introduce each of these research designs in the next set of slides and explain how they can help a researcher make a stronger causal claim.

Slide 77

Let's start with **repeated measures.** A repeated measures design requires that you measure the same thing at multiple points in time. At a minimum, this will be two points in time, which is the basic pre–post treatment design. Because cause precedes effect, we measure the outcome both before and after the law is implemented.

Slide 78

Perhaps not surprisingly, having more repeated measures is better than having fewer. This is because you might observe a time-lagged effect, or the law might have a curvilinear effect. Repeated measures over time and across jurisdictions offer the most powerful causal claims.

In the example provided here, you can see that if you had captured only two measurements, you might have assumed an effect when, in fact, there does not seem to be a uniformly upward or downward trend. Having just one observation before and after the intervention produces weak inference because any difference observed might reflect natural variation in the outcome over time.

Slide 79

Collecting dozens or hundreds of observations in a time series before and after a law is implemented can boost a team's confidence that an observed difference following implementation is attributable to the law.

A collection of observations over time can be viewed as one "time window," which is a single sample from an otherwise infinite time series. The larger the time window around a change in law, the easier it is to reliably assess the law's effects.

In this example, the first five O's represent observations at time points 1, 2, 3, 4, and 5. Then a law is passed, so you have an "interruption" in the observations. Observations 6, 7, 8, 9, and 10 are collected after the law has been implemented.

Slide 80

It is also essential to define a **time resolution,** which indicates whether you will measure your outcome data on a daily, weekly, monthly, or annual basis. A higher time resolution means more frequent measurements. To select an optimal time resolution, keep in mind some key considerations: speed, variation, scale, date, and seasonality.

Slide 81

Consider the **speed** at which a new law is expected to show effects. Laws change things slowly, and it might take a long time to produce an effect. You might miss an effect if you don't wait long enough or capture enough data over time. Or you might overestimate the magnitude of the effect if the initial passage of the law created a blip, but then over time, behavior reverted back to the way it was before the law was passed.

So, if effects are expected soon after a law's implementation, then using monthly observations rather than annual observations will make that effect easier to discern.

Slide 82

You'll also want to consider **variation** in the outcome at each time resolution. For example, if there is little to no variation in an outcome week by week, then monthly or annual measures might be more appropriate.

Slide 83

Another consideration is the **scale** of the measured outcome. For example, if you're trying to measure the number of vaccines given over time and you measure monthly, it might create a cloud-like scatter.

Slide 84

On the other hand, if you report yearly counts, you'll likely create a smoother curve.

Slide 85

The **date** when the law took effect is also an important consideration. For example, a mid-year effective date often does not work well with annual data. Let's say you can gather outcome data only at the annual level, and those data are released on September 1 each year. If you're studying a law that changed in May, it's unlikely that you will notice the effect of the law in the set of outcome data for that year.

Slide 86

Finally, consider the **seasonality** of the measured outcome. Behavior might be different in summer than in winter.

Slide 87

Now, let's discuss **functional forms of effects.** The functional form of effects relates to the shape of the line or curve of the outcome variable over time. With distracted driving laws, for example, the relationship between the amount of the fine and the reduction in deaths might be a linear or curvilinear relationship. If you increase the fine infinitely, does the reduction in deaths follow a linear rate of reduction? Or will there be a threshold at which increasing the amount of the fine no longer causes a reduction in deaths? That would suggest a curvilinear relationship.

Hypothesizing about the functional forms of legal effects brings up crucial questions that will shape the design of the study:

- Do you expect the effect to show up as soon as the law takes effect?
- Or do you expect a delay as enforcement or other implementation systems are developed and ramped up?
- Do you expect an anticipatory effect before the legal effective date, due to publicity and attention to the issue at the time the law is passed?
- Or maybe you expect the effect to emerge gradually, as various implementation systems change or as norms and behaviors shift?
- Do you expect the effect to be temporary, dissipating over time as organizations and individuals adapt to the new law in ways that maintain previous conditions?

These questions can help you and your researchers assess the type of functional form that you expect to recognize in the data. Each of these scenarios relates to the graphs on the next slide.

Slide 88

Think about the type of laws that might result in

- An abrupt change in behavior or outcomes,
- A temporary change in behavior or outcomes,
- A delayed change,
- An accelerated change,
- A gradual effect,
- A change in variability, or
- A partially decaying effect, in which the change in behavior gradually decreases.

Let's do the first one together. A law requiring restaurants to post calorie-count signage or face fines might result in an abrupt change. Such a regulation might almost immediately and dramatically increase the posting of signs inside restaurants.

This is a good time to pause and think through the different functional forms of effects. You can use examples from your own practice to work through the different functional forms.

Slide 89

Comparison jurisdictions are those that provide a basis for comparison. They differ from another jurisdiction in regard to the law or feature of law that you want to study. Two jurisdictions with different laws but similar socio-demographics and political systems might be well matched for a study.

Slide 90

When considering possible comparison sites, keep in mind that you will want to

- **Compare similar environments.** For example, if you are examining the effects of a new food policy, it would be critical to select comparison sites with similar socioeconomic and food environments.
- **Use multiple comparison sites to improve causal inference.** Having multiple comparison sites offers more opportunity for stronger causal claims.
- **Attribute differences between jurisdictions to the laws.** The goal is to be able to attribute any difference between the jurisdictions to the law.

And finally, you'll want to

- **Rule out alternative explanations.**

Slide 91

Now let's look at **comparison groups.** Comparison groups may refine the unit of analysis. If you're studying seat belt laws, for example, you might pick two comparison groups of all the males in one state and all the males in another state. Because males tend to have higher crash rates on average than females, this might be a way to simplify your study. This approach can minimize *noise,* or interfering data, that could be caused by other factors.

Selecting and including the most appropriate groups of individuals to compare is crucial, to ensure validity. For example, let's say a jurisdiction implements a regulation to reduce injuries in auto repair shops. After the law is passed, the jurisdiction experiences a reduction in injuries among auto repair workers but does not experience a reduction in injuries among workers in other settings. This finding would suggest that the regulation is effective.

Slide 92

Another way to strengthen your design is to use more than one outcome variable so that you can compare outcomes. This strategy helps to minimize threats to the internal validity of your study. For example, if you plan to examine the effects of motorcycle helmet laws, you might want to compare the rates of car and motorcycle fatality and injury rates. Assuming that a motorcycle helmet law does not affect car fatalities and injuries, this comparison would help rule out some other explanation for a reduction in motorcycle fatality and injury rates. Thus, you see that it is helpful to select a **comparison outcome** that is related to the law but not directly affected by it.

Slide 93

Another way to strengthen causal inference about a law's effects is to replicate the study across jurisdictions. **Replications** allow you to test the same effect more than once.

If similar effects are observed in each place where a similar law is implemented, we can infer that there is a strong causal relationship between the law and the overall effect.

Slide 94

Replication of studies can happen not only across multiple sites but also over time. If a law changes multiple times within one state or municipality, then you have multiple opportunities to test the effects of the changes.

In this slide, the top design indicates that outcome data were captured at three time points in a single jurisdiction before a law was enacted. Data for three additional time points were captured after the law was passed, and then the law was rescinded. Then, outcome data were captured at four subsequent time points. In other words, you have two sets of data for which there was *no law* in place.

The bottom design indicates that two observations of data were captured before a law was enacted. Then, two additional time points of data were collected before the law changed in some way. After that, two additional time points were collected. And so on. This process would give researchers multiple time points under two sets of legal landscapes.

For example, we know the causal effects of compulsory motorcycle helmet laws. Deaths decline abruptly when helmets become compulsory and abruptly return to higher levels when the law is rescinded.

Slide 95

Dose response is a clinical term, but it can apply to laws when one is discussing a group's level or amount of exposure to a law. Dose response is of interest when states have preempted localities from passing laws on topics like gun control, tobacco, and paid sick leave. Dose response is also of particular interest when states do not have a law, but local jurisdictions do. When a state passes a law and a city goes beyond the state law to legislate further, then the dose would differ for people who live in the state but don't live in the city with the stricter regulations.

Clean indoor air in Pennsylvania is a perfect example. The state does not have a clean indoor air law, but Philadelphia and Pittsburgh do, so the minute you leave either city, you can go into a bar or restaurant that allows smoking.

Be careful not to misappropriate a dose response. For example, if you looked at Pennsylvania state law and used average incidence of smoking-related illnesses in the state as a whole to compare against the outcomes in the two cities with a clean indoor air law, you would be "misappropriating dose" because you would not have accounted for how Philly and Pittsburgh affect the state's outcomes.

Slide 96

Let's walk through an example together. As you might know, in 1984 Congress passed a law requiring states to raise the minimum drinking age to 21 in order to receive federal highway funding. Some states were slow to raise their drinking age, while others moved quickly. If you wanted to examine the effects of raising the drinking age and use multiple design elements to strengthen your causal claims, you could do the following:

Select the states that had not yet raised the drinking age during the period of analysis. In this example, let's say the treatment states are Maine and Michigan, which initially kept the legal drinking age at 18 and older. The comparison states are New York and Pennsylvania, which raised the legal drinking age to 21 and older.

Slide 97

Within each state, you should stratify the drinkers by age group because you know the law might affect these three groups differently. Lumping them into one average might mask the effects. Thus, you would have these three groups:

- People who drink that are younger than 18
- People who drink that are ages 18-20
- People who drink that are 21 and over

The last group would be the comparison group. The drinking rates of the comparison group should not be affected by the law.

Slide 98

Within each state and within each age group, you will compare two sets of outcomes. You'll compare the rates of alcohol-related single-vehicle crashes that happen at night (in which the police report that drinking was a cause) against the rates of non-alcohol-related daytime crashes (in which the police do not report alcohol as a cause). This comparison allows you to tease out the effect of the law in a scenario where the number of overall crashes is going up for other reasons. It also allows you to examine the outcomes when you control for time of day.

Slide 99

In some cases, endogeneity becomes an issue. In plain English, **endogeneity** means you may have gotten the causal path wrong or there are alternative explanations for what you observe. As a result, the logic model – as well as the subsequent statistical model – does not properly capture the way causation works in the real world.

Let's use an example to illustrate what this might look like in practice. Consider this question: What led to an increase in seat belt use?

Slide 100

When we claim that seat belt laws cause people to wear seat belts, we assume that there is no alternative reason for the increase in wearing seat belts.

Slide 101

But suppose, for example, that right when seat belt laws were taking effect, carmakers began programming cars to make an annoying ringing sound until everyone in the car fastened their seat belts. In this model, we would have an endogeneity problem because we now have two potential and perhaps related causes for the observed increase in seat belt use. This problem is called *omitted variables endogeneity*. You can address this problem by adding the missing variable to the model, if you can measure it, and then trying to assess the extent to which the two factors influenced seat belt use individually and through their interaction.

Slide 102

Before moving to the next section, let's pause for a quick multiple-choice question. Which of the following design elements strengthens a quasi-experimental design?

A. Using many repeated measures

B. Using the highest time resolution possible

C. Comparing multiple jurisdictions

D. All of the above

Slide 103

If you selected D, that's correct! All three design elements strengthen a quasi-experimental design.

Slide 104

Once you have completed the planning phase, you can begin to operationalize the plan. We will now look at Phase 2: the operationalizing phase.

Slide 105

At this stage, you move from talking about concepts and features to identifying variables that you will use to "measure" the components in your logic model.

Like Phase 1, Phase 2 is iterative. We'll walk through each of these steps in the general process for operationalizing the plan:

- Building legal and outcome data
- Identifying an analytical plan
- Selecting variables
- Merging data for analysis

Phase 2 does not end until the data are clear and ready to be analyzed.

Slide 106

Wendy can begin operationalizing by conducting a full inventory of all the data sources that were identified in the planning phase. The type of legal epi study design will inform whether she needs legal or outcome data or both. Because she is studying the effects of a particular law on an outcome, she will need both legal data and outcome data.

Slide 107

After she's finished organizing her final data sources, Wendy will build or find an existing longitudinal, multi-jurisdictional dataset of distracted driving laws. One place she could find those data is www.LawAtlas.org.

Slide 108

Wendy's team would then need to figure out how to obtain state-level crash data – for example, the team could get data from the National Highway Traffic Safety Administration's website.

Slide 109

The epidemiologist or statistician on Wendy's team would download the two datasets and prepare them for analysis. This preparation would include reviewing codebooks, selecting variables, and paneling the data. *Paneling the data* means preparing data for longitudinal, time series analyses.

What is important to note here is the time that it will take to acquire data, clean data, and merge time-series data. This process takes a sophisticated data analyst one to three months. Getting to know every variable, its properties, and the range of values is a time-consuming but necessary step before moving on.

Slide 110

If you are planning a quantitative study, the operationalizing phase requires a team that includes a statistician, epidemiologist, or other quantitative methodologist. The study design and data sources drive the analytical plan.

Conducting an advanced legal epi study means you are testing causal links in the logic model. But there are steps that must happen before you can test causality.

If you have only observational data, you will be running descriptive and predictive analytics. **Descriptive statistics** are means, medians, modes, variances, and ranges. Regression techniques can be used in predictive and causal analyses, which can use the same statistical tests to address relationships. **Predictive analyses** are meant to predict whether one attribute is associated with another attribute, but they do not allow for testing causal links between one attribute and another. **Causal analyses** control for all potential confounding factors, so the research team can make a plausible statement like "attribute A causes attribute B." For example, with a sound research design and analytical plan, Wendy's team could say, "Evidence suggests that higher fines cause lower crash rates."

If your design allows for causal claims, then you can run more sophisticated types of multivariate regression.

Slide 111

The **difference-in-difference** model is the most common analysis for advanced legal epi studies. The name comes from the fact that you are estimating two types of differences simultaneously.

To understand what this would look like in practice, let's say Wendy's team is conducting a simple two-state study that compares state X and state Y to determine whether distracted driving laws affect traffic deaths. The number of traffic deaths is the outcome they are studying.

Slide 112

State X has a distracted driving law and state Y does not. Wendy's team measures the number of traffic deaths at two time points: one before the law was in place and one after the law was in place. All other things being equal between the two states, her team should observe a change in traffic deaths in state X, which has the distracted driving law, and they should observe no change in outcomes in the state that does not have a law on the books. The states should have a difference in difference – hence, the name of this approach.

Slide 113

It's common for a data project to begin with 30 to 60 variables and be whittled down to 5 to 10 variables of interest. Simplicity is best in regression models. Try to use the smallest number of variables needed to efficiently explain the variation in the outcome, unless the sample size is large enough to support more variables.

Slide 114

Preparing the data file for analysis is also a task for a quantitative methodologist. Identify early on which software package you will use to analyze the data; each software package requires a different data structure.

Slide 115

Now we're ready to talk about Phase 3, which focuses on analyzing data.

Slide 116

Once you've hit the analysis phase, any quantitative analyst can run through the standard univariate, bivariate, and multivariate techniques. These include descriptive statistics for each variable, crosstabs and t-tests for exploring bivariate relationships, and regression and causal analyses for predictive and causal inference.

Slide 117

Studying the impact of law is no different from studying the impact of any other intervention. Basic courses on quantitative analysis or statistics would provide guidance for the first two steps: descriptive and bivariate analyses.

Remember – plots, plots, plots! Examine the visual display of the bivariate relationships and the difference over time. Team up with an epidemiologist or quantitative methodologist to carry out the causal analyses and interpretations.

Slide 118

Every analysis begins with basic univariate descriptive statistics for the outcome variable and each input and mediator variable. The univariate descriptive statistics help identify missing data, outliers, and responses outside the range of plausible values. Data cleaning is an essential first step.

Univariate descriptive statistics are also referred to as *summary statistics*.

Descriptive statistics include at least one measure of central tendency (such as a mean, median, or mode) and at least one measure of spread (such as range, standard deviation, or variance).

The purpose of predictive and causal analyses is to find the right set of inputs and mediators that explain and predict the variance in an outcome. This purpose is why it is important to observe data over multiple places and time points. If there is little variance or spread in the outcome, there won't be anything to explain.

Slide 119

The second step of the analysis includes plots and bivariate tables between each input and mediator and the outcome. Plots will illuminate functional forms and relationships between inputs, mediators, and outcomes. In this graph, the percentage of traffic deaths in Colorado that were related to marijuana use is shown as a univariate plot of the outcome of interest. Note that there seems to be a steep uptick in marijuana-related traffic deaths after the recreational marijuana law was put in place, but not in the same year the law was passed. Perhaps drivers were more careful immediately after the law went into effect, and then the fear of penalty wore off. But can the upward tick be attributed to the legalization only? Logic models and hypothesis tests along with strong quasi-experimental designs and causal analytics could help answer this question.

You must follow descriptive plots with tests for association, such as chi-square tests for nominal or ordinal data and t-tests and F-tests for continuous outcomes.

Next, examine the correlation matrix to understand whether you have redundant information in the data. Can you omit variables that are eliciting the same or similar information? You can look at which variables appear to be individually correlated with the outcome.

Slide 120

We talked about analytical plans in Phase 2. Recall that many types of regression analyses appear in advanced legal epi studies. These include but are not limited to

- Random coefficient models such as difference in difference,
- Interrupted time series,
- Regression discontinuity,
- Hierarchical linear modelling, and
- Multivariate analysis of covariance.

We won't get into the nitty-gritty of how to conduct these analyses. Like Wendy, you can have an epidemiologist or a statistician help design and carry out the analysis.

Slide 121

Now you're ready to put all the pieces together. Reading through the results of your analyses is the fun part. At this time, you get to test the hypotheses you laid out in the beginning and start to understand whether there are statistically significant relationships in the model.

The first set of results often leads to a series of iterative runs of the models. Sometimes you might delete variables that are not significant, to make the model simpler. Once you have finished running the models and you are satisfied that you have answered the research questions you started with, you can move on to discussing the results.

Slide 122

There is always a discussion about whether a statistically significant relationship is of practical significance, and this discussion commonly starts with a conversation about the **size of the effect.** A relationship might have a p-value of less than .05, but the size of the effect might be so small that it makes the findings practically insignificant.

It's important to identify any strengths or weaknesses in the study design and the data. Readers of your study will want to know, for example, if you used proxy variables because you couldn't get a direct measurement of a feature or construct in the model. Be sure to discuss possible threats to validity and reliability of the study findings.

It's also important to highlight key results and identify next steps for future research. Was there a significant causal link between the presence of a law and an improvement in a health outcome? What do we still need to learn about the relationship between the law and the outcomes? Or do we know all there is to know?

Slide 123

Once you have your findings written up, you must decide how to share them. How will you disseminate the information to relevant stakeholders? Will you use peer-reviewed journals, policy briefs, research briefs, presentations, or infographics? To help answer these questions, consult with your communications and policy staff to identify key stakeholders, such as researchers, policymakers, policy analysts, public health practitioners, and other groups and organizations. Your staff can help you determine how best to communicate with each of these audiences and help you identify the types of materials that would be most useful to them.

Slide 124

Here's a review of what we've discussed. We first provided a quick recap of legal epidemiology. Next, we introduced the three phases of legal epi studies. In Phase 1, we described the steps in planning an advanced legal epi study. In Phase 2, we explored the advantages and limitations of various research designs; and in Phase 3, we focused on analyzing data.

Slide 125

The views expressed in this training, including in written training materials or publications, do not necessarily reflect the official policies of CDC or the US Department of Health and Human Services, nor does the mention of trade names, commercial practices, or organizations imply endorsement by the US government.

Like all trainings associated with the Public Health Law Academy, this training provides legal information but not legal advice, and no attorney-client relationship is created with the content developers of this training. Please contact an attorney in your state if you need legal advice.

Slide 126

Questions?

Thank you!