# Introduction to Causal Effects

## Confusion Over Causality

### Spurious Correlation

* Spurious correlation – unrelated variables that happen to be highly correlated with each other over a period of time.
* Often it’s not so clear. How about those cases where the relationship is claimed to be causal? How do we know it’s not spurious?
* Often headlines don’t use the term ‘cause’ or any of its variants, but the relationship gets interpreted causally. Need to read the details of the study to determine if the relationship was meant to be causal or not.
* Also, how sceptical a person is about the headline will depend on their point of view/prior beliefs.
* Need to move away from judging the form of the link based on prior beliefs and base it on the evidence; that is, how the study was designed, what statistical methods were used, what assumptions were made etc.

### Reverse Causality

* Reverse causality – when the direction of causality is unclear. For example, the relationship between green spaces and exercise: Are physically active people more likely to prioritise living near green space OR Does green space in urban environments cause people to exercise more?
* To sort this out, would need to carefully examine the temporal relationships between the variables.

### Clearing Up the Confusion

* Causal inference/causal modelling attempts to do so by proposing:
  + Formal definitions of causal effects.
  + Assumptions required for identifying causal effects from data.
  + Rules about what variables need to be controlled for.
  + Sensitivity analyses to determine the impact of violations of the assumptions on the conclusions.

### History

* Own area of statistical research since the 1970s.
* Causal diagrams – generally thought to be very helpful in describing the hypothesised relationships between the variables of interest. Theory has also been developed to identify the variables you need to control for to identify the causal effects.
* Propensity scores – major contribution in estimating causal effects.
* Time-dependent confounding – treatments/exposure vary over time and exposures at one time affect lots of other variables at future times leading to a never-ending feedback loop. So how do you estimate the joint causal effect of treatment over time?
* Optimal dynamic treatment strategies – what treatment should be provided given a subject’s characteristics; not just – is treatment A better than treatment B?
* Targeted learning – ML approach based on semi-parametric theory and high dimensional data.

### Going Forward

* Focus on causal inference from observational studies and natural experiments.
* Causal inference requires making some untestable assumptions (i.e. causal assumptions)
  + i.e. can’t check these are true from data – which is why sensitivity analysis is important.

## Potential Outcomes and Counterfactuals

### Treatments and Outcomes

* Treatment and exposure can be used interchangeably.
* Most of the time the treatment will be considered as a binary variable where 1 means the subject received the treatment and 0 otherwise.
* Outcomes can be binary or continuous.
  + Observed outcomes – outcomes seen in reality
  + Potential outcomes – outcomes we would observe under each possible treatment option. i.e. theorising about what would happen under each treatment option before conducting the study.
* Notation: is the outcome that would be observed if treatment was set to
  + Each subject has potential outcomes
* For example, suppose the treatment is the influenza vaccine and the outcome is the time until the subject gets the flu.
  + is the time until the subject would get the flu if they received the vaccine
  + is the time until the subject would get the flu if they did not receive the vaccine.
* For example, the treatment is regional or general anaesthesia for surgery. The outcome is surgery complications.
  + Interested in whether the type of anaesthesia lead to higher/lower risk of surgery complications.
  + is 1 if there are surgery complications and 0 otherwise if given regional anaesthesia.
  + is 1 if there are surgery complications and 0 otherwise if given general anaesthesia.

### Counterfactuals

* Counterfactual outcomes are ones that would have been observed had the treatment been different.
  + If the treatment was , the the counterfactual outcome is .
  + If the treatment was , the the counterfactual outcome is
* For example, did the influenza vaccine prevent the subject from getting the flu?
  + What actually happened:
    - The subject received the vaccine and did not get sick.
    - The actual exposure was .
    - The observed outcome was .
  + What may have happened:
    - Had the subject NOT received the vaccine, would he/she have gotten sick?
    - The counterfactual exposure is .
    - The counterfactual outcome is .

### Potential Outcomes and Counterfactuals

* Before the treatment decision is made, any outcome is a potential outcome: and
* After the study, there is an observed outcome , and counterfactual outcome
  + (Notation assumes a binary treatment so and ).
* Counterfactual outcomes are typically assumed to be the same as potential outcomes.
  + Hence, the terms potential and counterfactual outcomes are used interchangeably.

## Hypothetical Interventions

### Intervention

* Typically defined as something that can be manipulated.
* Difficult to consider the causal effects of a variable that can’t be manipulated (even if only in the hypothetical).
* Common to assume that there are no hidden versions of treatment.
* For example, if we were interested in the causal effect of BMI on health outcomes, we would have a problem because:
  + There are many ways that a subject could achieve a BMI of a particular value.
  + These methods may also be associated with different outcomes.
  + Weight is not directly manipulable, so it may be better to consider the causal effects of interventions that aim to manipulate weight.

### Immutable Variables

* i.e. things like race, gender or age.
* Makes causal inference difficult as you can’t manipulate these variables.

### Manipulable vs Non-Manipulable

* One way around this is to consider a proxy – a related variable that can be manipulated.
  + For example, studying the discrimination of hiring practices with respect to race. You can’t change the race of a subject, but you could change the name on an otherwise identical resume and observe if that makes a difference.

### Causal Effects

* Focus on treatments/exposures that could be thought of as interventions i.e. at a hypothetical level at a minimum.
  + Treatments that we could imagine being randomised (manipulated) in a hypothetical trial (even if it wouldn’t be ethical to do so).
* There are causal effects of these immutable/non-manipulable variables, but they don’t fit as cleanly into the potential outcomes framework (which will be used to identify causal effects in this course).
* Focus on hypothetical interventions because:
  + Their meaning is well-defined.
  + Potentially actionable (very important – i.e. you then end up with outcomes you can do something with).

### What Are Causal Effects?

* had a causal effect on if differs from i.e.
  + A causal effect occurs when the potential outcomes and are not equal to each other; i.e. that the potential outcome under treatment differs from the potential outcome under no treatment.
  + For example
    - The outcome is that the headache is gone one hour from now and the treatment is taking ibuprofen.
    - The reality is that only one treatment is observed: “I took ibuprofen and the headache is gone, hence the medicine worked.” – but this isn’t proper causal reasoning. All they told you was .
    - But if you hadn’t taken ibuprofen, would the headache have gone anyway i.e. what is the counterfactual: .

### Fundamental Problem of Causal Inference

* Can only observe one potential outcome for each subject.
* BUT, with certain assumptions, can estimate population (average) causal effects.
  + Can never know a unit level causal effect – i.e. what would have happened to that particular subject under the alternative treatment.
  + Can know the average outcome if all subjects underwent the treatment vs if no one did.
  + Are able to say something about how well exposure and treatment works in general at a population level NOT at an individual level.

## Causal Effects

### Average Causal Effect

* Population of interest – everyone you’re interested in.
* Ideally would like to see two hypothetical worlds:
  + World 1: Everyone gets treatment .
  + Word 2: Everyone gets treatment .
* The important thing is that the population in the two worlds are exactly the same; just different treatments.
* If we could observe both populations simultaneously, can collect the outcome data and take the average value. The difference between the two is the Average Causal Effect for the population.
* In reality, this can’t happen.
* More formally : Average value of if everyone was treated with minus the average value of if everyone was treated with .
  + If is binary, this is a risk difference because the mean of a binary variable is just a probability or risk.

### Example 1

* Population of interest: people undergoing hip fracture surgery.
* Outcome: surgery complications.
* Treatment: Anaesthesia type (regional vs general).
* Suppose average causal effect = . This means that the probability of surgery complications is lower by 0.1 if given regional anaesthesia compared with general anaesthesia.
  + In plain English: the probability of surgery complications is lower by 0.1 if given regional anaesthesia compared to general anaesthesia OR if 1,000 people were going to have hip fracture surgery, we would expect 100 fewer to experience surgery complications under regional anaesthesia compared with general anaesthesia.

### Example 2

* Population of interest: people with hypertension.
* Outcome: systolic blood pressure.
* Treatment: Hypertensive medication or no treatment.
* Suppose average causal effect = . This means that in the population of hypertensive patients, the average systolic blood pressure is 20 mm Hg lower if they took anti-hypertensive medication compared to if they did not.

### Conditioning On, Versus Setting, Treatment

* In general,
  + Conditioning in this context isn’t the same as the usual statistical sense. The main idea is that the average causal effect (the mean of potential outcomes) is generally not going to be the same as the expected value of the observed given minus the expected value of the observed given .
  + Restricting to the subpopulation of subjects who actually had the treatment .
  + The expected value of among the subpopulation of subjects who have .
  + But the subpopulation might differ from the population as a whole in important ways e.g. people at higher risk for flu might be more likely to choose to get a flu shot, which means that we’re taking the for a higher-risk population.
  + This is different than because this represents the outcome if everyone in the whole population received the treatment (not restricting to a subpopulation).
  + Setting treatment = manipulating in the potential outcome situation.
  + Conditioning on = restricting to subpopulations.

### Real World

* Some subjects received the treatment and some didn’t.
* Could take the average of both subpopulations, which is the average difference in the outcome between the subpopulations defined by the treatment group. But these subpopulations may differ in fundamental ways, so a treatment effect hasn’t been isolated.

### Conditioning Versus Setting

* In general the mean of treated subjects minus the mean of untreated subjects is NOT a causal effect because two different subpopulations are being compared.
  + And these two subpopulations might differ in important ways independent of the treatment.
* IS a causal effect because it’s comparing what would happen if the same population were all treated with versus i.e. the only thing that’s different is the treatment and hence a treatment effect has been isolated.

### Other Causal Effects

* Could be interested in other causal effects:
  + Causal relative risk:
  + Causal effect of treatment on the treated:
    - Restricting to a subpopulation (the treated group), but are contrasting potential outcomes for that group so are still in causal effect territory.
    - Interested in how well treatment works among treated people – useful because there may be some subpopulations that may never be interested in the treatment.
  + Average causal effect in the subpopulation with covariate :
    - Heterogeneity treatment effects where there might be some subpopulation defined by variable .
    - Want to know the causal effect in this subpopulation.

### Causal Effect of Treatment on the Treated

* Only interested in the treated subpopulation.
* So now in world 1 vs world 2, the same subpopulation is given and , so it’s a causal effect.
* Wish to imagine what would have happened if we had not treated the subpopulation in World 1, which is provided in World 2.

### Challenge

* But we return to the fundamental problem of causal inference, which is that we only observe a single treatment and outcome for each subject.
* How do we use observed data to link observed outcomes to potential outcomes? i.e. How do I estimate causal effects from observational data?
  + Need to make assumptions.

## Causal Assumptions

### Identifiability

* Identifiability of causal effects requires making some untestable assumptions – causal assumptions.
  + Not able to estimate/test these assumptions with data.
* Most common:
  + Stable Unit Treatment Value Assumption (SUTVA)
  + Consistency
  + Ignorability
  + Positivity
* Assume the observed data will include the outcome, treatment and a set of covariates.

### SUTVA

* Really is two assumptions:
  + No interference – subjects do not interfere with each other; that is, the treatment assignment of one subject doesn’t affect the outcome of another subject.
    - Subjects don’t interact with each other.
    - Also called spillover or contagion.
    - There are approaches that can handle interference, but these aren’t covered in this course.
  + One version of treatment
    - Important as if there’s multiple variants of the treatment, it becomes difficult to define a causal effect as well as causing other problems.
  + SUTVA allows us to write the potential outcome for the ith subject in terms of only that subject’s treatments instead of having to account for the treatments in the rest of the population.
  + In many cases a reasonable assumptions and makes things a lot simpler.

### Consistency

* Directly linking the outcome and the observed data: The potential outcome under treatment  
   is equal to the observed outcome if the actual treatment received is .
* The observed outcome is equal to the potential outcome if treatment is equal to a for any possible treatment:

### Ignorability

* Possibly the most important assumption and sometimes known as the ‘no unmeasured confounders’ assumption.
* The basic idea is that treatment assignment is assumed to be independent from potential outcomes conditional on the pre-treatment variables/covariates; i.e. if the covariates are correct then treatment is effectively randomly assigned.
* Among subjects with the same values for the confounding variables, treatment if effectively randomly assigned.
  + Random in the sense that it’s independent of the potential outcomes (and may not be random in some other sense).
  + Treatment assignment itself becomes ignorable.
* For example:
  + is a single variable (age) that can take the values ‘younger’ or ‘older’.
  + Older people are more likely to get treatment , but also more likely to get a hip fracture irrespective of treatment.
  + Age is hence related to treatment and the risk of the outcome, so treatment is not randomly assigned and and are not independent of .
  + BUT within levels of X, treatment is effectively randomly assigned.
* The question then becomes what covariates need to be included to make this assumption of ignorability reasonable.

### Positivity

* The idea that all subjects had some chance of getting either treatment, and that’s conditional on the covariates.
* At every level of and for every treatment, every subject had a non-zero probability of getting treatment OR treatment isn’t deterministic as a function of .
* Using the previous example, it would be a violation of the assumption if all older subjects were treated, but not a violation if all subjects could get treated but older subjects were more likely to get treated.
* Recall that this assumption is required because we need data where we can learn about what would happen under either treatment scenario, which can only occur if there is some chance of subjects receiving either treatment option within every level of .
  + Also need it to be not deterministic as otherwise there’s some situations in which certain subjects could never receive a particular treatment.
  + Then need to exclude them from the study and only make inferences about the population that have some chance of being treated.
* Variability in treatment assignment important for identification.

### Observed Data and Potential Outcomes

* Can put these assumptions together to identify causal effects.
* involves only observed data.
  + Expected value of among the subpopulation of subjects who have the treatment equal to and covariates .
  + No potential outcomes: are all observed.
* by the consistency assumption (as long as the latter holds).
  + Which means we go from something involving only observed data to something involving potential outcomes.
* by the ignorability assumption.
  + Ignorability means we can drop the conditioning on treatment .
  + Conditioning on isn’t providing us any additional information about the mean of the potential outcome because as long as you condition on , it’s randomly assigned.
* If we want a marginal causal effect, we need to average over the distribution of .
  + One that doesn’t involve conditioning on .

## Stratification

* One way in which causal effects can be estimated or identified.
* Stratify on important variables and then average over the distribution of those variables – also known as standardisation.

### Conditioning and Marginalising

* Combined is also known as standardisation.
* Recall that the average causal effect is defined as: , which doesn’t involve any covariates. But need to condition on to be able to link the observed outcome to the potential outcome.
  + To get rid of the covariates, we just average over the distribution of .
  + Suppose there’s a single categorical variable, then:
  + This is the expected value of the potential outcome, which means we can contrast it and get causal effects.
  + = average/sum of all levels of x
  + Expected value of the potential outcome is equal to the expected value of the outcome in the subpopulation of interest averaged over the distribution of the covariates.
* Known as standardisation where conditioning means stratifying and marginalising means averaging over.
* End up with a standardised mean, which happens to be the same as the average potential outcome.

### Standardisation

* Involves stratifying and then averaging i.e. obtaining a treatment effect within each stratum and then pooling across the stratum weighting by the probability/size of each stratum.
* From data, treatment effects are estimated by computing the means under each treatment within each stratum and then pooling across the stratum.
* For example, consider a study comparing two diabetes treatments.
  + Outcome = major adverse cardiac event (MACE).
  + Challenges:
    - Users of one treatment were more likely to have had past use of another drug.
    - Patients with past use of those other drugs are at a higher risk for a MACE in general.
  + Key idea:
    - Compute the rate of MACE for the two treatments in two subpopulations:
      * Patients with no prior other drug use.
      * Patients with prior other drug use.
    - Take the weighted average where weights are based on the proportion of people in each subpopulation.
    - Causal effect if, within the levels of the prior other drug use variable, it is reasonable to assume diabetic treatment assignment was random (ignorability given prior other drug use).
      * E.g the main thing clinicians based their treatment decision on was prior drug use.
  + Raw, unstratified data

|  |  |  |  |
| --- | --- | --- | --- |
|  | **MACE = 1** | **MACE = 0** | **Total** |
| **Treatment A** | 350 | 3,650 | 4,000 |
| **Treatment B** | 500 | 6,500 | 7,000 |
| **Total** | 750 | 10,250 | 11,000 |

* + - Probability of MACE given Treatment A = 350 / 4000 = 0.088
    - Probability of MACE given Treatment B = 500 / 7000 = 0.071.
    - Based on the raw data, subjects given Treatment A have a worse outcome. But don’t know if that’s due to Treatment A being less effective than Treatment B or Treatment B was preferentially assigned to subjects that were sicker anyway.
  + Stratify on the X variable: Prior other drug use. i.e. create two 2x2 tables instead of one.
    - Prior other drug use = NO

|  |  |  |  |
| --- | --- | --- | --- |
|  | **MACE = 1** | **MACE = 0** | **Total** |
| **Treatment A** | 50 | 950 | 1,000 |
| **Treatment B** | 200 | 3,800 | 4,000 |
| **Total** | 250 | 4,750 | 5,000 |

* + - Prior other drug use = YES

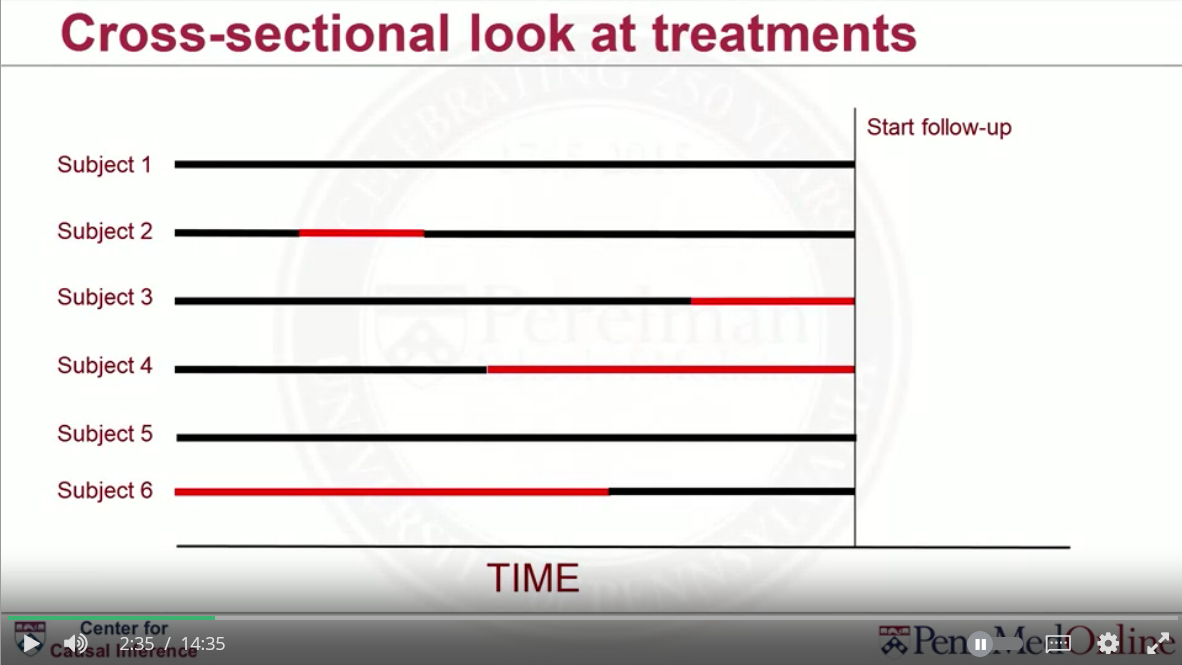
|  |  |  |  |
| --- | --- | --- | --- |
|  | **MACE = 1** | **MACE = 0** | **Total** |
| **Treatment A** | 300 | 2,700 | 3,000 |
| **Treatment B** | 300 | 2,700 | 3,000 |
| **Total** | 600 | 5,400 | 6,000 |

* + Treatment A recipients are more likely to have prior drug use.
    - Treatment A total | Prior other drug use = 3,000 (YES) vs 1,000 (NO)
    - Treatment B total | Prior other drug use = 3,000 (YES) vs 4,000 (NO)
  + Subjects with prior other drug use are at higher risk of MACE irrespective of treatment.
    - MACE | Prior other drug use = 600 / 6000 = 0.1 (YES)
    - MCAE | Prior other drug use = 250 / 4000 = 0.0625 (NO)
  + When prior other drug use = NO there is no difference in the risk of MACE between the two treatments:
    - Treatment A = 50 / 1000 = 0.05
    - Treatment B = 200 / 4000 = 0.05
  + When prior other drug = YES there is also no difference in the risk of MACE between the two treatments:
    - Treatment A = Treatment B = 300 / 3000 = 0.1
  + Collectively can conclude that neither treatment carries a higher risk of MACE once prior other drug use is accounted for.
  + Now we’re going to look at the mean potential outcome for Treatment A (previously were just comparing the rates of the outcome in the different subpopulations) as we want to marginalise (to get an expected value of a potential outcome not conditional on X).
    - Expected value of Y had everyone been assigned Treatment A:
  + Calculate the expected value of Y (probability of MACE) among Treatment A subjects at each level of X and then take a weighted average based on the size of the corresponding populations.
  + Once we marginalise, the mean of the expected value for both treatments is exactly the same i.e. the potential outcome is the same if all subjects were given Treatment A or all subjects were given Treatment B.
  + This approach can become problematic very quickly if there are many X variables required to reasonably assume ignorability.
    - For example, multiple combinations of X variables with no data, so no way to calculate a mean and average.
  + Solutions in the course focus on IPTW and propensity score methods for observational data and instrumental variable analysis for natural experiments where there is a variable that could be considered a randomiser.

## Incident User and Active Comparator Designs

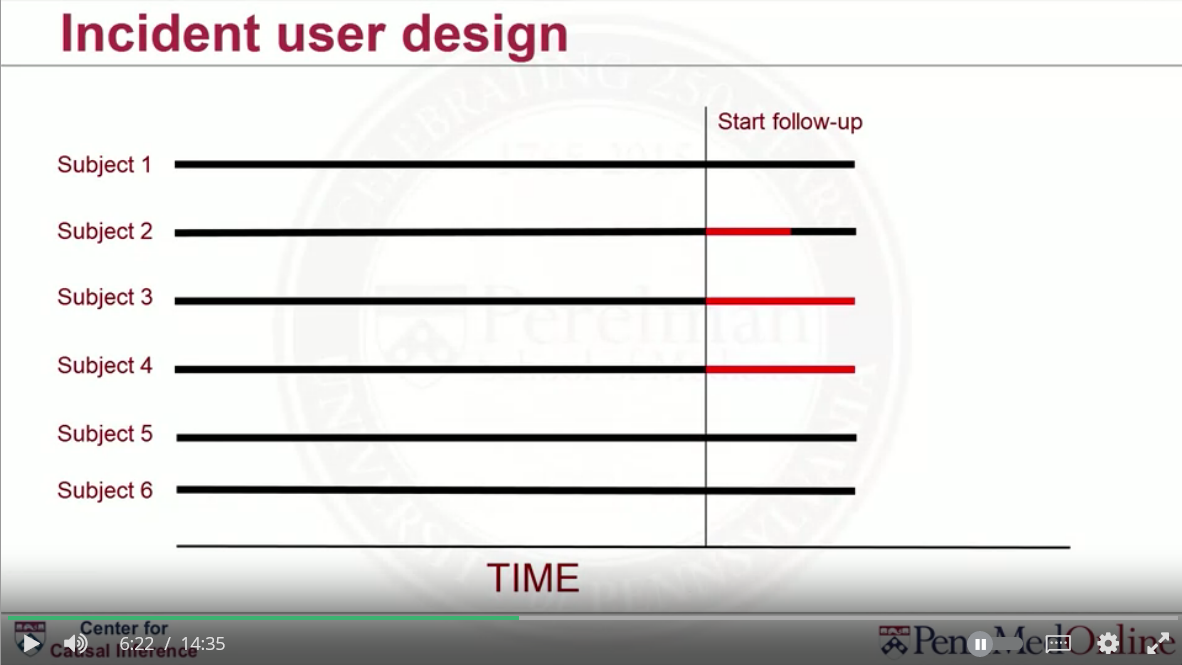
### Cross-Sectional Look at Treatments

* Suppose we are interested in whether yoga effects blood pressure:
  + At any given time, some people regularly practice yoga whilst others do not.
  + Those who are currently practicing may only have just started or may have been doing so for a while.
  + Those who are not practicing may have done so in the past.
    - Why did these people stop?
    - This a type of selection bias that is very difficult to control for, especially when considering a cross-sectional snapshot instead of the entire history.
    - Also there may be some lingering treatment effects for people who practiced for a long time and then stopped.



## Incident User Design

* Also known as new user design.
* Aims to get around these problems by restricting the treated population to subjects newly initiating treatment.
  + In the yoga example, this design excludes subjects who had practiced yoga in the past.



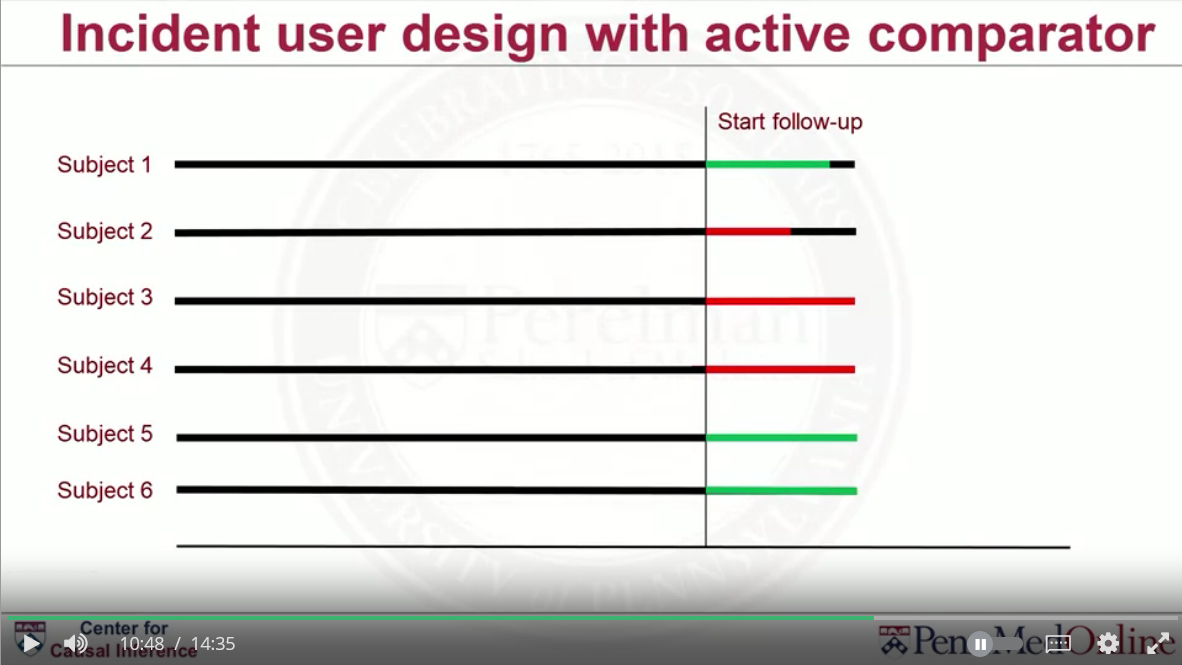
* The causal question is slightly modified to the causal effect of initiation, but this results in a cleaner problem.
* If the comparison group is no treatment, not always clear when follow-up should start; for example, could match on characteristics.
* Alternative is to use any active comparator, as opposed to no treatment, which will be defined here as an inactive comparator.
  + Much cleaner.
  + For example, some other fitness option like Zumba fitness.

### Active Comparator Design

* Tend to involve much less confounding.
  + People who practice yoga or Zumba are probably more alike than people who don’t exercise at all.
* But the causal question is changed again and more narrow.
  + i.e. the effect of yoga vs the effect of Zumba on blood pressure compared to the effect of yoga to everything else/no exercise.
* Very much depends on the research question as to which design is more suitable/appropriate.

### Incident User Design with Active Comparator

* Very useful for causal inference problems.



* Prior to t=0, all subjects were taking none of the treatments (and hadn’t in the past).
* At t=0, all subjects start one of the possible treatments.
* In theory, all subjects are somewhat similar because they’re all starting a treatment and researchers can control for the confounding variables that are measured up to t =0.

### Other Considerations

* Sometimes it’s not possible to implement an incident user design e.g. the causal effect of air pollution.
* Sometimes no treatment (unexposed) is the comparison group of interest so an active comparator isn’t applicable.
* This course is focussed on a treatment-initiation kind of design/situation, but there are causal methods that can handle time-variant treatments (the causal effect of treatments over time) e.g. treat over a full year vs over part of the year.