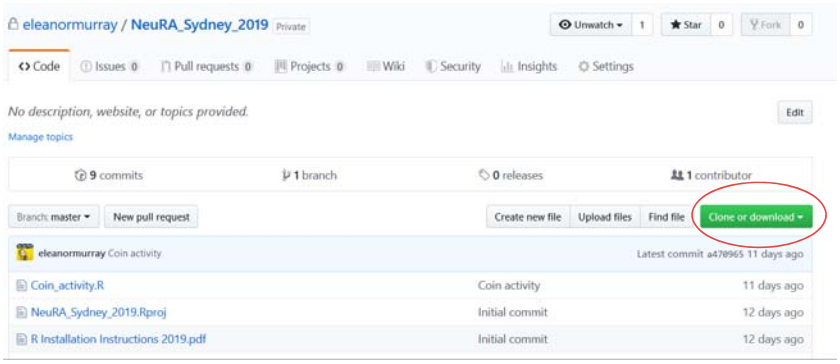
All materials available at:
https://github.com/eleanormurray/NeuRA_Sydney_2019/



Getting the most out of pragmatic trials – beyond the intention-to-treat effect

Eleanor Murray, ScD
Department of Epidemiology
BU School of Public Health

NeuRA Sydney
Nov 20-21, 2019

@EpiEllie

Day 2: Estimating causal effects from pragmatic trials with survival outcomes: a coding practicum

Schedule Today

1. Overview
2. Estimating intention-to-treat effects
3. Estimating per-protocol effects

Part VII: Practicum overview

Why are we here and what are we doing?

This practicum applies what we learned yesterday estimating causal effects in a trial with a survival outcome

We'll estimate the intention-to-treat effect (a point exposure) and the per-protocol effect (a static sustained exposure)

The practicum materials were jointly developed by Ellen Caniglia and Lucia Petito

The case study: Coronary Drug Project (CDP)

NHLBI funded trial of lipid-lowering medications in men with a history of heart attacks

- Enrollment began in 1966, and follow-up ended in 1975
- 5 active treatments versus placebo

Coronary drug project research group *JAMA 1975*

The case study: Coronary Drug Project (CDP)

Protocol	Description
Eligibility criteria	Men with a history of a myocardial infarction in previous 3 months 30-64 years old
Treatment arms	5 lipid-influencing drugs vs. placebo
Follow-up	Randomization until 5 years, LTFU, or death
Outcome	5-year mortality risk
Causal contrasts of interest	Intention-to-treat effect Effect of good adherence to trial protocol versus poor adherence, stratified by randomization arm

Coronary drug project research group *JAMA 1975*

A little about our data

The datasets on github are simulated versions of the Coronary Drug Project trial

The data are in long format, which means we have 1 observation per person per visit

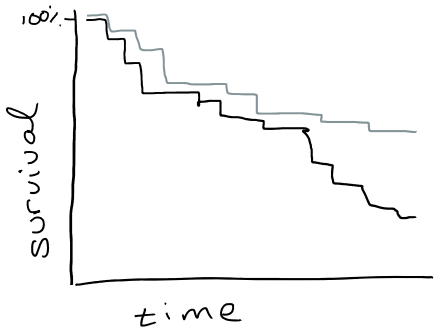
ID	time	Z	L_0	L_t	C_t	Y_t
1	0	1	0	0	0	0
1	1	1	0	1	0	0
1	2	1	0	1	0	1
2	0	0	1	1	0	0
2	1	0	1	0	0	0
2	2	0	1	0	1	.
3	0	1	1	1	0	0
3	1	1	1	0	0	0
3	2	1	1	0	0	0

We have 15 time-varying covariates, and 1 baseline-only variable

Variable	Notation	Definition	Values
simID	id	ID variable	Range: 0 to 4042
rand	Z	Randomization group	0: placebo; 1: clofibrate
visit	t	Visit number: visits occur quarterly	Range: 0 to 14
death	Y_t	Death recorded at visit t	0: alive; 1: died $T \in [t, t+1)$
adhpre0bin	A_{-1}	Adherence to placebo during pre-randomization run-in period	0: adherence $\geq$ 80%; 1: adherence $<$ 80%
adhr_b	A_0	Adherence to placebo at baseline (recorded at a special visit 2 weeks after randomization)	0: adherence $\geq$ 80%; 1: adherence $<$ 80%
adhr	A_t	Adherence to assigned treatment at visit t	0: adherence $\geq$ 80%; 1: adherence $<$ 80%
mi_bin	L	Myocardial infarction at baseline	0: 1 or 2; 1: $\geq$2
AntiHyp	L	Antihypotensive medication use at visit t	0: no; 1: yes
AnyQQS	L	ECG findings: Q/QS pattern at visit t	0: no; 1: yes
AnySTDep	L	ECG findings: ST depression at visit t	0: no; 1: yes
AP	L	Angina pectoris at visit t	0: no; 1: yes
CardioM	L	Cardiomegaly at visit t	0: no; 1: yes
CHF	L	Coronary heart failure at visit t	0: no; 1: yes
DIUR	L	Diuretic use at visit t	0: no; 1: yes
FVEB	L	ECG finding: Frequent ventricular beats at visit t	0: no; 1: yes
HiHeart	L	High heart rate at visit t	0: $<$70 bpm; 1: $\geq$ 70 bpm
HiSerChol	L	High serum cholesterol at visit t	0: $<$250; 1: $\geq$ 250
HiSerTrigly	L	High serum triglyesterol at visit t	0: $<$5.0; 1: $\geq$ 5.0
IC	L	Intermittent claudication at visit t	0: no; 1: yes
NIHA	L	New York Heart Association class at visit t	0: no limitations; 1: any limitation
OralHyp	L	Oral hypoglycemic agents use at visit t	0: no; 1: yes
VCD	L	ECG finding: Ventricular conduction defect at visit t	0: no; 1: yes

Survival analysis basics

Survival analysis allows us to compare the **time-to-event** rather than just the number of events



The first challenge: administrative censoring

Let's define T as the time to death

- T=1 for subjects who die in month 1
- T=2 for subjects who die in month 2, etc.
- T is unknown for subjects who are alive at end of follow-up

This is called administrative censoring: time of death T is unknown for subjects who have not died by end of follow-up

Administrative censoring is different from loss to follow-up

Some people will drop out of our study. For these

- T is unknown after some point during follow-up

This is called loss to follow-up. We often need to deal with this in our studies, but for the workshop we have data with no loss to follow-up.

Some survival analysis definitions

Survival probability: $\Pr[T>k]$

Risk (cumulative incidence): $\Pr[T\leq k]$

- All deaths between baseline and time *k*
- Compared to the number of individuals at baseline

Discrete hazard: $\Pr[T=k+1|T>k]$

- Recent deaths occurring during interval *k* to *k+1*
- Compared to the number individuals alive (and under follow-up) at *k*

Causal inference for survival analysis is difficult because

- Time-varying confounding-exposure feedback
- Time zero & immortal time bias
- Exposures can be hard to define well

These problems are less likely with a randomized trial, but common in observational studies

Reminder, why do we bother with intention-to-treat effects?

Randomization ensures no confounding at baseline for treatment assignment

Randomization also ensures positivity for treatment assignment

Randomization is a well-defined intervention

So, intention-to-treat analyses often give unbiased estimates of intention-to-treat effects

Exercise: Let's draw an intention-to-treat DAG

Some notation you could use:

Z = randomization

L = confounders

A = adherence

Y = outcome

C = loss to follow-up

Next, let's draw a per-protocol effect DAG

Finally, let's draw an intention-to-treat DAG with loss to follow-up

Part VIII: Estimating intention-to-treat effects

Exercise 3.1: Data exploration

3.1 Data exploration

Before we estimate the intention-to-treat effect, let's get familiar with the data. Make sure you have the workshop materials downloaded and unzipped. In your preferred software, open the code:

- For R, double click on the '.Rproj' file in the 'CausalSurvivalWorkshop_2019' folder to open the workspace.
- For SAS, navigate to the 'SAS' folder and double click the '.sas' file.
- For Stata, navigate to the 'Stata' folder and double click the '.do' file to set Stata's working directory to the workshop folder.

Run the code in **Code Section 0**. This ensures that all required libraries are installed (for R), the working directory is correctly assigned to the workshop folder (R & SAS), and the data is loaded into the software (R, SAS, Stata).

Once you have your coding environment set up, run the code in **Code Section 1** and answer the following questions:

Question 1

How many person-visits are in this dataset?

Page 8

1. Open the code in your preferred language
2. Work through Code Sections 0 & 1
3. Answer Questions 1-5

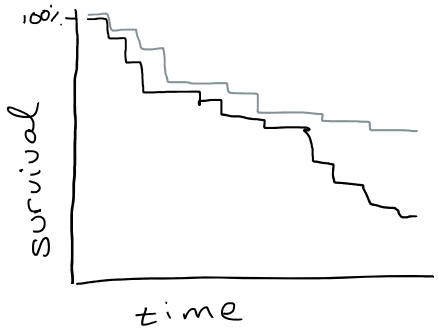
Non-parametric survival estimates

Since we have no loss to follow-up, we can estimate the causal effect of randomization on survival without making any additional assumptions

We can do this by comparing the Kaplan-Meier survival curves

Reminder: Survival curves

Kaplan-Meier curves look like staircases



Kaplan-Meier survival

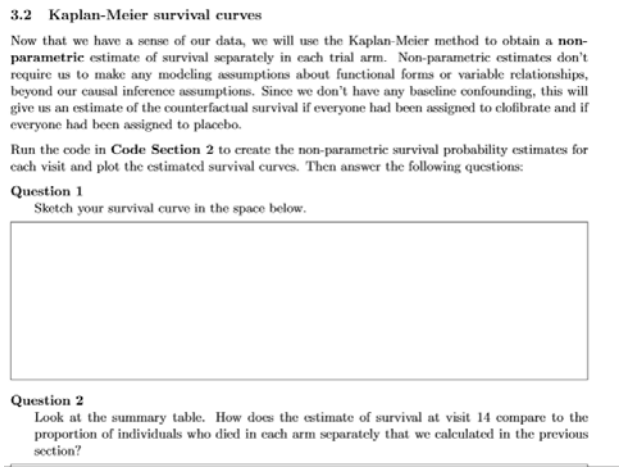
Kaplan-Meier estimates survival at *k* as the product of conditional probabilities of having survived each interval between 0 and *k*

The conditional probability at each time is one minus the discrete hazard

For example, for $k=2$

$$\Pr[D_2=0] = \Pr[D_1=0]\Pr[D_2=0|D_1=0]$$

Exercise 3.2: Kaplan-Meier Survival Curves



3.2 Kaplan-Meier survival curves

Now that we have a sense of our data, we will use the Kaplan-Meier method to obtain a **non-parametric** estimate of survival separately in each trial arm. Non-parametric estimates don't require us to make any modeling assumptions about functional forms or variable relationships, beyond our causal inference assumptions. Since we don't have any baseline confounding, this will give us an estimate of the counterfactual survival if everyone had been assigned to clofibrate and if everyone had been assigned to placebo.

Run the code in **Code Section 2** to create the non-parametric survival probability estimates for each visit and plot the estimated survival curves. Then answer the following questions:

Question 1

Sketch your survival curve in the space below.

Question 2

Look at the summary table. How does the estimate of survival at visit 14 compare to the proportion of individuals who died in each arm separately that we calculated in the previous section?

1. Work through Code Section 2
2. Answer Questions 1-3

Semi-parametric ITT estimate

We can easily estimate the intention-to-treat effect on the hazard ratio scale if we're willing to make one additional assumption:

- The hazard ratio is constant over follow-up

Now we can use Cox proportional hazards regression

$$\lambda(t \mid Z, L_0) = \lambda_0(t)\exp(\alpha_1 Z + \alpha_2 L_0)$$

Parametric ITT estimate

If we're willing to make one more assumption, we can estimate the hazard ratio and the survival, plus the risk difference and risk ratio

- The baseline hazard has a specified functional form
- (here we assume quadratic)

Now we can use pooled logistic regression

$$logit[\Pr(Y_t = 1 \mid Z, L_0)] = \beta_{0,t} + \beta_1 Z + \beta_2 L_0 = \beta_0^* + \beta_1 Z + \beta_2 L_0 + \beta_3 t + \beta_4 t^2$$

MEJ10

Exercise 3.3.1: Unadjusted Intention-to-Treat Effect

Run the code in **Code Section 3a**, and fill out the first 3 rows of Table ??.

Exercise 2 results

		Cox proportional hazards	Pooled logistic regression
Unadjusted	Coef		
	SE		
	HR		
Adjusted	Coef		
	SE		
	HR		

Question 1
Do the results from the *unadjusted* Cox proportional hazards model and the *unadjusted* pooled logistic regression model match?

Question 2
What is the causal interpretation of these unadjusted hazard ratios?

Question 3
What assumptions are we making to give this hazard ratio a causal interpretation?

1. Run Code Section 3a
2. Fill in the first 3 rows of the table
3. Answer questions 1-3

Slide 34

MEJ10 Add questions
Murray, Eleanor J, 11/17/2019

Baseline covariate adjustment

Both semi-parametric and parametric survival methods allow adjustment for baseline covariates to estimate conditional hazard ratios

Exercise 3.3.2: Baseline-adjusted Intention-to-Treat Effect

Run the code in **Code Section 3b**, complete the last 3 rows of Table ??, and answer the following questions.

Question 1
Do the results from the *adjusted* Cox proportional hazards model and the *adjusted* pooled logistic regression model match?

Question 2
Do the results from the *adjusted* models match the results from the *unadjusted* models?

Question 3
What is the causal interpretation of these adjusted hazard ratios?

1. Run Code Section 3b
2. Fill in the last 3 rows of the table
3. Answer questions 1-5

Murray. 2019. Causal Survival Analysis

Maintaining interpretability

Conditional hazard ratios are hard to interpret

We would prefer to know the average hazard ratio, or even better the average survival and risks

We can estimate these from our pooled logistic regression model using standardization

- (Aka the g-formula)

A trick to standardizing without calculating probabilities

id	rand	sex	age	race	death
1	0	0	65	1	0
2	1	0	58	1	1
3	0	1	67	1	1
4	1	0	54	0	0
5	0	1	48	0	0

1. Original data

Exercise 3.4 Standardizing over baseline covariates to estimate the average intention-to-treat effect

Run the code in **Code Section 4**, fill in the table below, and answer the following questions.

visit (t)	$\hat{S}^{z=0}(t)$	$\hat{S}^{z=1}(t)$	RD	HR	CIR
0				- - -	- - -
1				- - -	- - -
2				- - -	- - -
3				- - -	- - -
4				- - -	- - -
5				- - -	- - -
6				- - -	- - -
7				- - -	- - -
8				- - -	- - -
9				- - -	- - -
10				- - -	- - -
11				- - -	- - -

1. Run Code Section 4
2. Fill in the table
3. Answer questions 1-4

Part IX: Estimating per-protocol effects

What is a per-protocol analysis?

Common approaches

- censor when non-adherent, don't adjust for confounding
- add adherence to regression model, adjust only for baseline confounders

Common ≠ correct!

Per-protocol analyses tell us

how did trial outcomes differ between those who did adhere to, or recieved, assignment A and those who did adhere to, or receive, assignment B?

Per-protocol analyses in the literature

Approach	Description

Per-protocol analyses in the literature

Approach	Description
1. "Modified ITT"	▪ censor never initiators

Per-protocol analyses in the literature

Approach	Description
1. "Modified ITT"	▪ censor never initiators
2. "Per-protocol population"	▪ censor if never initiate, cross-over, or discontinuation

Per-protocol analyses in the literature

Approach	Description
1. "Modified ITT"	▪ censor never initiators
2. "Per-protocol population"	▪ censor if never initiate, cross-over, or discontinuation
3. "As-treated"	▪ allow cross-over ▪ censor non-initiators or discontinuers

Methods 1 to 3:
Censor without adjustment

Per-protocol analyses in the literature

Approach	Description
1. "Modified ITT"	▪ censor never initiators
2. "Per-protocol population"	▪ censor if never initiate, cross-over, or discontinuation
3. "As-treated"	▪ allow cross-over ▪ censor non-initiators or discontinuers
4. Adherence adjustment	▪ include adherence in model for outcome model

Methods 1 to 3: Censor without adjustment

Method 4: Adjustment for baseline confounding only

Potential per-protocol analyses

Approach	Description
1. "Modified ITT"	▪ censor never initiators
2. "Per-protocol population"	▪ censor if never initiate, cross-over, or discontinuation
3. "As-treated"	▪ allow cross-over ▪ censor non-initiators or discontinuers
4. Adherence adjustment	▪ include adherence in model for outcome model
5. Instrumental variables, aka "contamination-adjusted ITT"	▪ compare outcome by trial arm, and correct using adherence by trial arm

Methods 1 to 3: Censor without adjustment

Method 4: Adjustment for baseline confounding only

Effects are different from analyses

Per-protocol effect tells us

> "how would trial outcomes differ if everyone adhered to assignment A versus if everyone adhered to assignment B"

Better per-protocol analyses

Approach	Description
Per-protocol effect estimation	▪ censor if deviate from protocol or include adherence in outcome model ▪ adjust for censoring or time-varying confounding
5. Instrumental variables, ~~aka "contamination-adjusted ITT"~~	▪ compare outcome by trial arm, and correct using adherence by trial arm

Exercise 4.2: Data cleaning for the per-protocol effect

Run the code in **Code Section 5** and answer the following questions.

Question 1

How many individuals in the placebo arm are adherent at baseline? How many individuals in the treatment arm are adherent at baseline?

Question 2

Comment on the Kaplan-Meier curves you generated. After visually inspecting the data, do the curves look similar?

1. Run Code Section 5
2. Answer questions 1-2

Inverse probability weighting

Step 1: identify adherent person-time

Step 2: build inverse probability weights for adherence and fit separately in each trial arm

Step 3: censor when non-adherent

Step 4: fit an IP-weighted model for the outcome given trial arm, baseline covariates

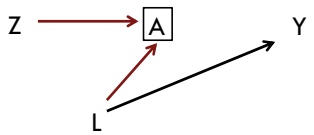
Step 5: generate estimated survival & risks standardized over baseline covariates

Inverse probability of adherence weights

- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[A_j=a_j|Z,\bar{L}_j,\bar{A}_{j-1}]}$
- $SW_t = \prod_{j=0}^{t} \frac{Pr[A_j=a_j|Z,\bar{A}_{j-1}]}{\Pr[A_j=a_j|Z,\bar{L}_j,\bar{A}_{j-1}]}$
- At each time, each person receives a weight inversely proportional to the probability of the adherence pattern they have, conditional on randomization, time-varying covariates, and adherence history

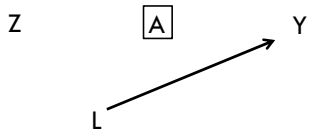
Adjusting for non-adherence

- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[A_j=a_j|Z,\bar{L}_j]}$



Adjusting for non-adherence

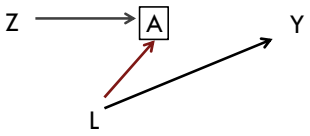
- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[A_j = a_j | Z, \bar{L}_j]}$



Non-stabilized weights create a pseudo-population with no adherence problems!

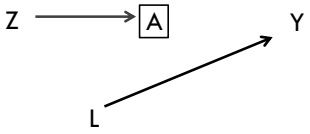
Adjusting for non-adherence

- $SW_t = \prod_{j=0}^{t} \frac{Pr[A_j = a_j | Z]}{\Pr[A_j = a_j | Z, \bar{L}_j]}$

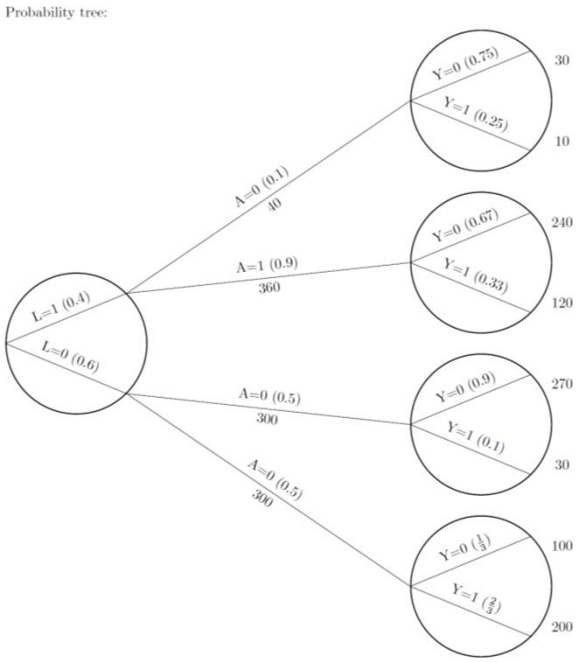


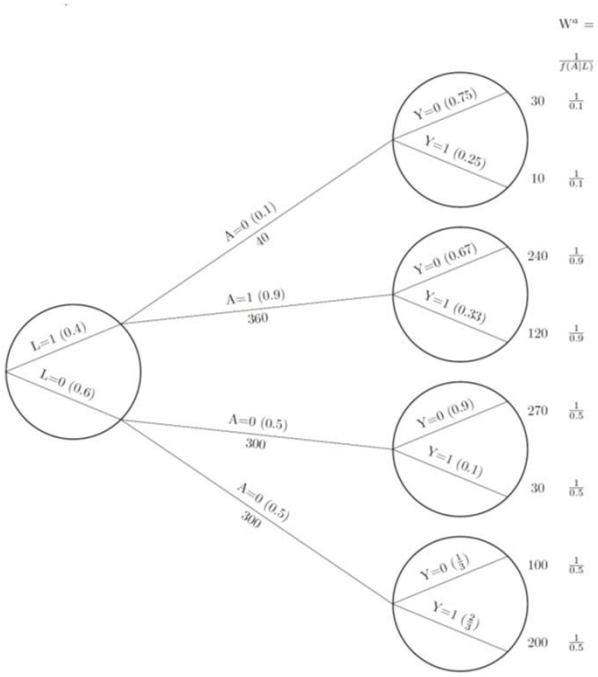
Adjusting for non-adherence

- $SW_t = \prod_{j=0}^{t} \frac{Pr[A_j=a_j|Z]}{\Pr[A_j=a_j|Z,\bar{L}_j]}$



Stabilized weights create a pseudo-population with selection but no selection bias!





Estimating the per-protocol effect with IPW

Step 1: identify adherent person-time

Step 2: build inverse probability weights for adherence and fit separately in each trial arm

Step 3: censor when non-adherent

Step 4: fit an IP-weighted model for the outcome given trial arm, baseline covariates

Step 5: generate estimated survival & risks standardized over baseline covariates

Exercise 4.3.1: Estimate inverse probability of adherence weights

Run the code in **Code Section 6**, fill in the table, and answer the following questions.

	Unstabilized	Stabilized	Truncated Stabilized (99^{th} percentile)
Mean (SD)			
Range			
Median (IQR)			
99th Percentile			

Question 1

What was the mean of your stabilized weights after truncation? What did you expect it to be? Did stabilizing the weights (compared to unstabilized) change the mean of the weights? How did truncating change the mean and range of the weights?

1. Run Code Section 6
2. Fill in the table
3. Answer question 1

Wrap-up

What we learned

For point exposures:

- Non-parametric survival curves
- Unadjusted hazard ratios from semi-parametric and parametric models
- Conditional hazard ratios from semi-parametric and parametric models
- Average hazard ratios, survival curves, risk differences, and cumulative incidence ratios from parametric survival models

What we learned

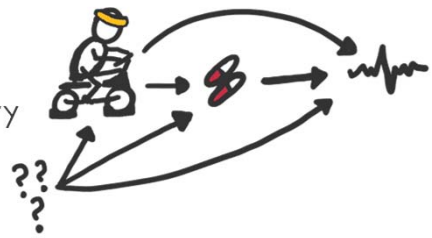
For static sustained exposures:

- Non-parametric survival curves
- Unadjusted hazard ratios from semi-parametric and parametric models
- Conditional hazard ratios from semi-parametric and parametric models
- Average hazard ratios, survival curves, risk differences, and cumulative incidence ratios from parametric survival models

What about observational studies?

Observational studies have confounding

If we have a point exposure, we need to worry about baseline confounding



If we have a sustained exposure, we need to worry about baseline and time-varying confounding

We've learned how to handle both already!

Some more complicated scenarios you may encounter:

- Loss to follow-up and non-adherence – address this by multiplying inverse probability weights
- Dynamic sustained strategies – these typically require unstabilized weights
- Grace periods – whenever someone is following multiple strategies, clone them, and when you know which strategy they're following censor them & adjust with IPW
- Competing events – think carefully about the causal effect of interest

Where to get more information

Some references:

- Proposed pragmatic trial guidelines: https://www.hsph.harvard.edu/causal/pragmatictrials/
- Patient-centered causal effects: Murray et al. 2018. J Clin Epi 103:10-21.
- Choosing a causal effect: Hernan & Scharfstein. 2018, Ann Intern Med; 168(7):515-6.
- Per-protocol effect estimation: Hernan & Robins. 2017, NEJM 377:14; Lodi et al, 2016. AIDS; 30(17):2659-63.
- Placebo arm adherence analyses: Murray & Hernan. 2018, Trials 19:158; Murray & Hernan. 2016, Clin Trials 13(4): 372-8.
- G-methods: Causal Inference, Hernan & Robins. Available online at: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

Contact me:

@EpiEllie
ejmurray@bu.edu
https://github.com/eleanormurray

THANK YOU FOR COMING