Segment 2: Randomized Studies

Section 05: "Breaking" Randomized Experiments

Randomized Assignment Doesn't Guarantee Everything

- (Randomized) treatment assignment guarantees balance (on average) on the basis of *pre-treatment* features of the units
- What happens post-randomization:
  1. May not be under the experimenter control
  2. May be a affected by the randomization
  3. May not be "random" in the sense that it may relate to the potential outcomes
  4. This could "break" the randomization

Noncompliance

- Typically, only the **assignment** to treatment is randomized
- But whether units actually **comply** with treatment is not always controlled
- Even when *treatment assigned* is random, *treatment received* may not be
  - E.g., if only the most health-conscious participants actually *receive* the supplements
  - It is typically *not* appropriate to assume whether a unit complies or not is random (ignorable)
- Randomization preserves ability to estimate the causal effect of **treatment assignment**
  - May or may not correspond to the causal effect of **treatment receipt**

Examples of Treatment Noncompliance

- Hypothetical dietary study
  - People assigned to take supplements may choose not to
  - People assigned control may decide to take supplements anyway
- Military service, randomized draft lottery
  - People drafted to serve in military do not serve
  - People who are not drafted enlist anyway
- Job training program
  - Those randomized to a job training program do not complete all of the training sessions
- Maxillofacial surgery
  - Surgeon "overrules" a randomized surgical treatment because of features of the patient's injury
- Pregnancy trial
  -

What to do with treatment noncompliance?

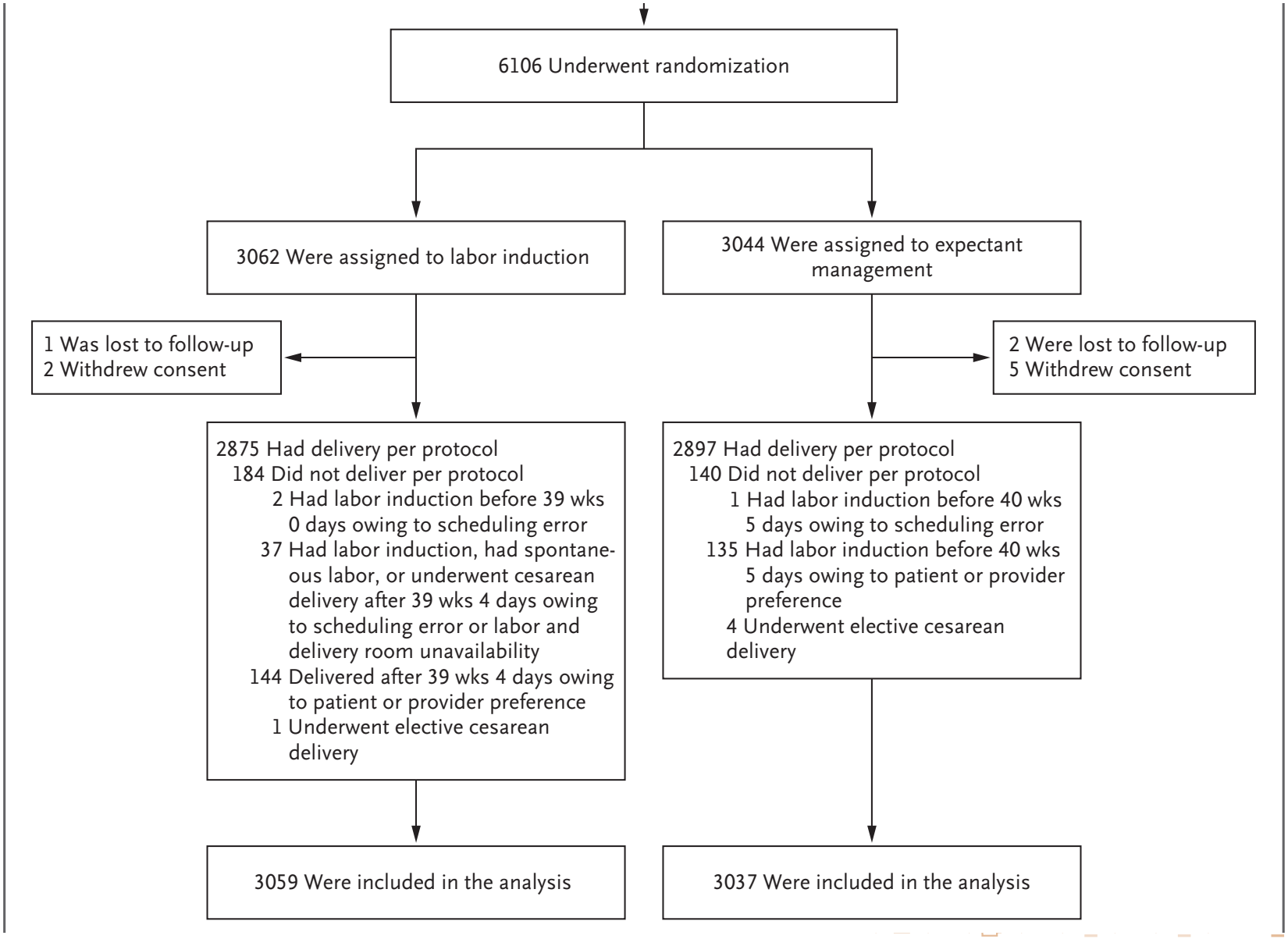
Some options, there are others

Some options:

1. *Intent to Treat Analysis*
   - Analyze units on the basis of their *randomized* treatment, regardless of whether they actually received it
   - $\bar{Y}|Z = 1 - \bar{Y}|Z = 0$
   - Preserve randomization....but for the *assignment*, not the actual *receipt* of treatment
   - Is a causal effect....but maybe not the one we care about
2. *As Treated Analysis*
   - Analyze units on the basis of which treatment was actually *received*
   - $\bar{Y}|D = 1 - \bar{Y}|D = 0$
   - Does not enjoy the benefits of randomization
   - Is not a causal effect
3. *Per Protocol Analysis*
   - Analyze the subset of patients who actually received whichever treatment was randomly assigned

Example: ARRIVE Pregnancy Trial

Cropped Figure 1



Example: Oral and Maxillofacial Surgery

Do the Benefits of Rigid Internal Fixation of Mandible Fractures Justify the Added Costs? Results From a Randomized Controlled Trial

*Vivek Shetty, DDS, Dr MedDent,**
Kathryn Atchison, DDS, MPH,† Richard Leathers, DDS,‡
Edward Black, DDS,§ Cory Zigler, MS,‖ and
Thomas R. Belin, PhD¶

Purpose: Owing to its putative advantages over conventional maxillomandibular fixation (MMF), open-reduction and rigid internal fixation (ORIF) is used frequently to treat mandible fractures, particularly in noncompliant patients. The resource-intensive nature of ORIF, the large variation in its use, and the lack of systematic studies substantiating ORIF attributed benefits compel a randomized controlled investigation comparing ORIF to MMF treatment. The objective of this study was to determine whether ORIF provides better clinical and functional outcomes than MMF in noncomplying type of patients with a similar range of mandible fracture severity.

Patients and Methods: From a total of 336 patients who sought treatment for mandible fractures, 142 patients with moderately severe mandible fractures were assigned randomly to receive MMF or ORIF and followed prospectively for 12 months. A variety of clinician and patient-reported measures were used to assess outcomes at the 1, 6, and 12 months follow-up visits. These measures included clinician-reported number of surgical complications, patient-reported number of complaints, as well as cumulative costs of

In Terms of Potential Outcomes

- $Z \in \{0, 1\}$ denote randomized treatment assignment
  - $Z \perp\!\!\!\perp Y^0, Y^1 | X$
- $D \in \{0, 1\}$ denote which treatment is actually *received*
  - Noncompliance: $Z_i \neq D_i$
- **Key Point:** D is *posttreatment*!
  - Like an intermediate outcome
- Warrants the potential outcomes notation
  - D_i^0: Treatment that would be received if assigned $Z_i = 0$
  - D_i^1: Treatment that would be received if assigned $Z_i = 1$
- Can define potential outcomes in terms of Z and/or D

$$Z \perp\!\!\!\perp Y^0, Y^1 | X \not\Rightarrow D \perp\!\!\!\perp Y^0, Y^1 | X$$

Key Idea: Need to Think About (D_i^0, D_i^1)

Z_i	X_i	D_i^0	D_i^1	Y_i^0	Y_i^1
0	$\star$	$\star$	?	$\star$	?
0	$\star$	$\star$	?	$\star$	?
0	$\star$	$\star$	?	$\star$	?
0	$\star$	$\star$	?	$\star$	?
...					
1	$\star$	?	$\star$	?	$\star$
1	$\star$	?	$\star$	?	$\star$
1	$\star$	?	$\star$	?	$\star$
1	$\star$	?	$\star$	?	$\star$

Example: Hypothetical Dietary Experiment

With treatment noncompliance

Table: Observed Data from the Hypothetical Dietary Experiment, **Idealized Assignment**

Unit, i	Female, x_{1i}	Age, x_{2i}	Treatment Z_i	Potential D_i^c	Potential D_i^t	Potential Y_i^c	Potential Y_i^t
Audrey	1	40	0	0	1	140	135
Anna	1	40	1	0	0	140	135
Bob	0	50	0	0	0	150	140
Bill	0	50	1	0	0	150	140
Caitlin	1	60	0	0	1	160	155
Cara	1	60	1	0	1	160	155
Dave	0	70	0	1	0	170	160
Doug	0	70	1	0	1	170	160

Missing Data

Two broad categories of missing data:

1. Missing covariate values
   - That is, missing *pre-treatment* quantities
   - E.g., Unit features
   - Not *that* big of a deal in terms of estimating causal effects

2. Missing outcomes
   - That is, missing *post-treatment* quantities
   - E.g., missing response values
   - Can be a *very* big deal in terms of estimating causal effects

Missing Data: Pretreatment Covariates

- Whether data are missing pre-treatment is not related to randomization
  - Missing data should still be *balanced* by the design
- Ignoring covariate with missing data won't induce bias
  - E.g., X_i is missing for some $i \Rightarrow$ don't adjust the analysis for X_i
  - Efficiency implications
- Dropping units with missing data won't induce bias
  - But focuses analysis on a subset of units
  - May sacrifice *external validity*
- In both cases, may increase standard error of effect estimates

Missing Data: Outcomes

- Missingness could depend on the treatment assignment
  - That is, treatment actually affects which units have missing Y
  - A unit's missing outcome would have been observed under the alternative treatment
- Ignoring units with missing outcome values can induce bias
  - Randomization guarantees balance (on average) of potential outcomes
  - Treatment-induced missingness could "break" the would-be balanced distribution of potential outcomes
- Focusing only on units with complete data could negate the benefits of randomization
- Analogy with noncompliance in that both are examples of dealing with *posttreatment variables* that are actually potential outcomes
- **Don't adjust for posttreatment variables**

Example: Hypothetical Dietary Experiment

With missing outcome data

Table: Observed Data from the Hypothetical Dietary Experiment, **Idealized Assignment**

Unit, i	Female, x_{1i}	Age, x_{2i}	Treatment Z_i	Missing? M_i^c	Missing? M_i^t	Potential Y_i^c	Potential Y_i^t
Audrey	1	40	0	0	1	140	135
Anna	1	40	1	0	0	140	135
Bob	0	50	0	0	0	150	140
Bill	0	50	1	0	0	150	140
Caitlin	1	60	0	0	1	160	155
Cara	1	60	1	0	1	160	155
Dave	0	70	0	1	0	170	160
Doug	0	70	1	0	1	170	160

Example: Growth Hormone Study
(from *Missing Data in Longitudinal Studies* by Daniels and Hogan)

Setting: Randomized trial examining the effect of recombinant human growth hormone (rhGH) on muscle strength in the elderly

- Study arms: Exercise plus Placebo (EP) vs. Exercise plus rhGH (EG)
  - Completely Randomized Design
- Outcome: quadriceps strength (QS) measured at baseline, 6 months, and 12 months follow up
- Goal: Compare mean QS at 12 months follow up

Missing Data in the GH Example

Treatment	s	n_s	Month 0	Month 6	Month 12
EG	1	12	58 (26)		
	2	4	57 (15)	68 (26)	
	3	22	78 (24)	90 (32)	88 (32)
	All	38	69 (25)	87 (32)	88 (32)
EP	1	7	65 (32)		
	2	2	87 (52)	86 (51)	
	3	31	65 (24)	81 (25)	73 (21)
	All	40	66 (26)	82 (26)	73 (21)

General Precaution: Beware Posttreatment Variables

- Randomization can sometimes be a starting point
- In practice, many things can "go wrong"
- Generally speaking, things that "go wrong" *after* randomization can be problematic
  - *After* randomization $\Rightarrow$ (possibly) affected by randomized assignment
  - No longer random
  - "Break" the ignorability assumption that holds with respect to Z by design