500 Class 11

https://thomaselove.github.io/500-2025/

2025-04-03

Agenda for the Slides

- Highlights from Lehr et al. (2019/2020)
- Questions and Answers about the Project

(if time permits)

- Dealing with Time-Varying Covariates and Related Issues
- Augmenting Sensitivity Analysis

Section 1

Lehr et al. (2019)

The impact of change in definition of increased-risk donors on survival after lung transplant

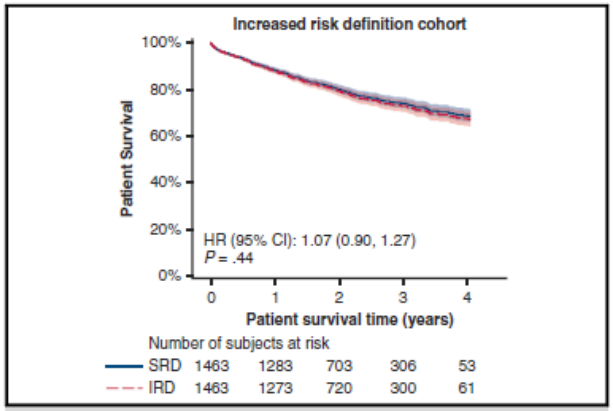
Carli J. Lehr, MD, MS,[a] Rocio Lopez, MS,[b] Susana Arrigain, MA,[b] Jesse Schold, PhD, MStat, MEd,[b] Christine Koval, MD,[a] and Maryam Valapour, MD, MPP[a]

ABSTRACT

Objectives: To study the impact of using the US Public Health Service broadened definition of "increased-risk" donors (2013) in comparison with "high-risk" (1994) and standard infectious risk donors on lung transplant recipient outcomes.

Methods: Patients who underwent lung transplant between January 1, 2006, and May 31, 2017, in the Scientific Registry of Transplant Recipients were divided into 2 cohorts, recipients of: (1) high-risk donors: January 1, 2006, to October 1, 2013, and (2) increased-risk donors: January 1, 2014, to May 31, 2017, and compared with matched recipients who received standard-risk donors. Risks for acute rejection, patient, and graft survival using propensity score matched cohorts, multivariable logistic, and Cox models were examined.

Results: In total, 18,490 lung transplant recipients were analyzed with 36% transplanted during the increased-risk donor definition period. The proportion of donors classified as nonstandard infectious risk increased with the definition



Patient survival by donor risk category. Survival was similar for recipients of standard-risk (*SRD*) and increased-risk donor (*IRD*) organs.

CENTRAL MESSAGE

The donor risk definition update

Lehr et al. (2019) Central Message

CENTRAL MESSAGE

The donor risk definition update in 2013 increased the number of donors classified as nonstandard risk. The use of increased-risk donors expands the donor pool without impacting survival.

Propensity Score (from Lehr 2019)

A logistic regression model was used to estimate the propensity score; having an HRD/IRD was modeled as the outcome with all recipient and donor characteristics (see next slide) as independent variables.

- A greedy matching algorithm was used to select the best match from the same definition cohort for each recipient; all HRD/IRD recipients were matched.
- The standardized differences in all covariates before and after matching were evaluated to assess matching success. (See Figure 2 in two slides.)

Outcomes

The primary outcome was posttransplant patient survival. Secondary outcomes included graft survival and acute rejection (treated) within 1 year of transplant. Outcomes were assessed using the matched cohort. Patient and graft survival were evaluated using Cox regression with a robust sandwich covariance matrix estimate to account for intracluster dependence due to matching.

Recipient and Donor Characteristics

Recipient characteristics included in the propensity score model were age, sex, body mass index, race, primary source of payment for transplant, median income in past 12 months by ZIP code, education, history of cigarette use, diabetes, Lung Allocation Score (LAS), lung disease group, Karnofsky functional group, life support, previous lung surgery, previous transplants for other organs, most recent serum creatinine, total bilirubin before transplant, forced expiratory volume in 1 second (FEV1) %, infection requiring intravenous drug therapy within 2 weeks before transplant, transplant year/definition of high risk, time on waiting list, total ischemic time, and procedure type.

In addition, the following donor characteristics were also included: age, sex, body mass index, race, ABO group, alcohol use, history or recent use of cigarettes, cocaine, other drugs, death by stroke, death by non-beating heart.

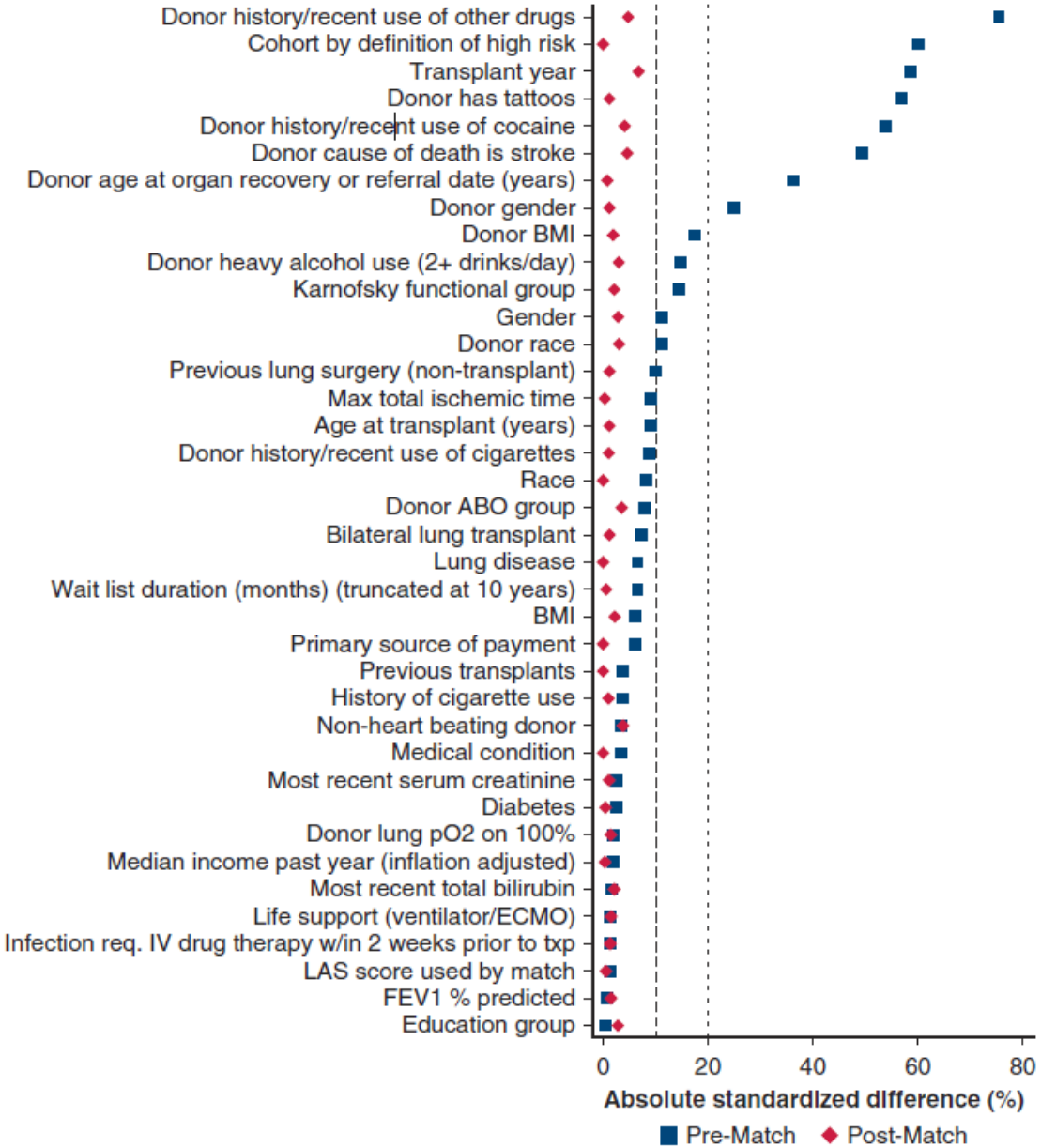


FIGURE 2. Standardized differences before and after matching. *Blue squares* indicate the prematch characteristics (original cohort, n = 18,490), and *red diamonds* indicate the postmatch comparison (matched cohort, n = 4774). The *dashed* and *dotted lines* represent absolute standardized differences of 10% and 20%. *BMI*, Body mass index; *max*, maximum; pO_2, partial pressure of oxygen; *ECMO*, extracorporeal membrane oxygenation; *IV*, intravenous; *LAS*, Lung Allocation Score; *FEV1*, forced expiratory volume in 1 second.

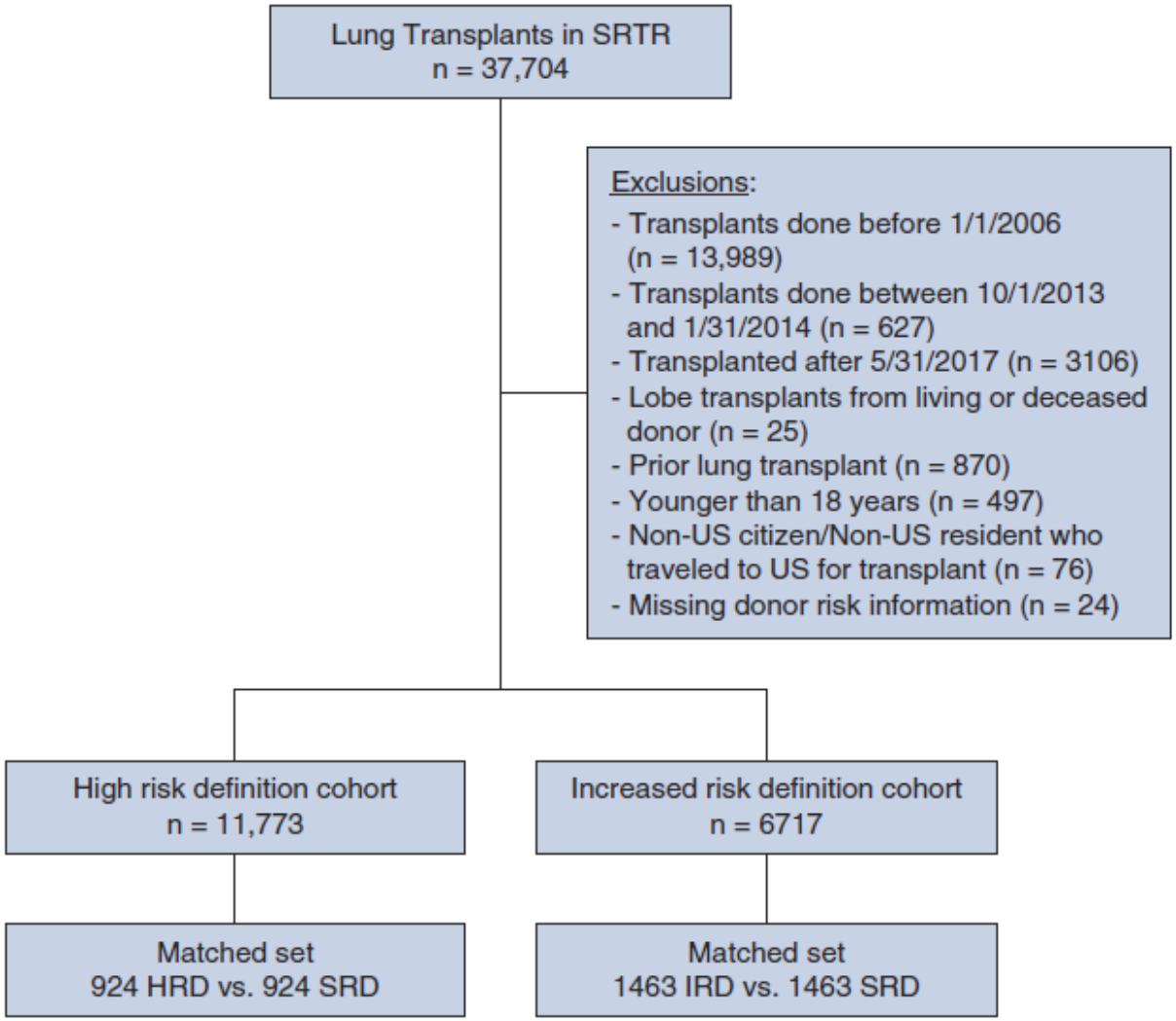


FIGURE 1. Flow diagram of the study population All lung transplant recipients between January 1, 2006, and May 31, 2017, were selected for analysis and divided into the high-risk donor cohort (January 1, 2006, to October 1, 2013) and increased-risk donor cohort (February 1, 2014, to May 31, 2017). The time frame when either classification could be used (October 1, 2013, to January 31, 2014) was excluded from analysis. *SRTR*, Scientific Registry of Transplant Recipients; *US*, United States; *HRD*, high-risk donor; *SRD*, standard-risk donor; *IRD*, increased-risk donor.

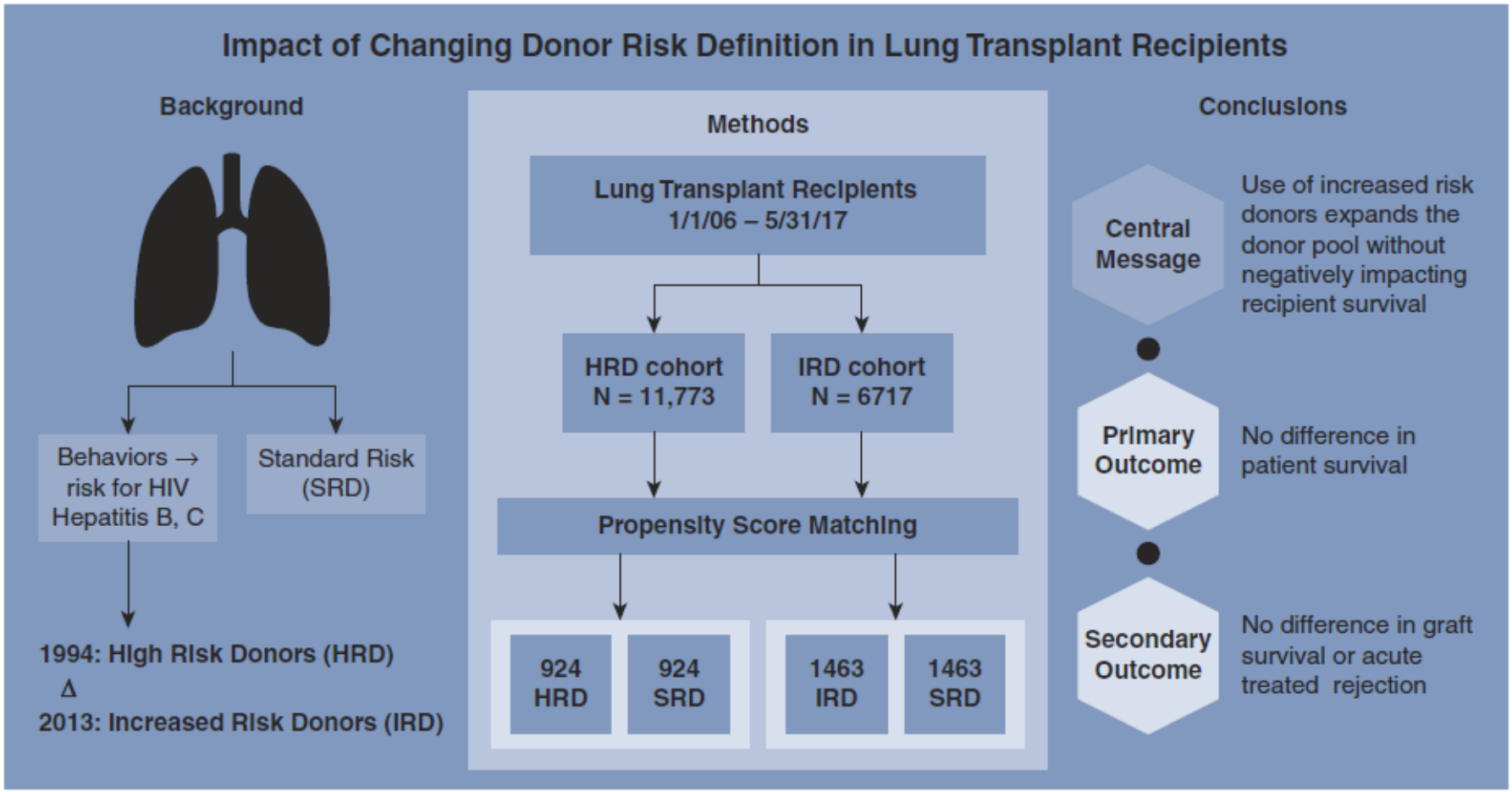


FIGURE 6. Background, study methods, and study conclusions. Donors at elevated risk for human immunodeficiency virus (*HIV*), hepatitis B virus (HBV), and hepatitis C virus (HCV) defined by Public Health Service (PHS) as high-risk donor (*HRD*) in 1994, which was changed to a more inclusive definition of increased-risk donor (*IRD*) in 2013 (*left panel*). Cohort selection and propensity matching are summarized in the *middle panel*. The *right panel* highlights the central message, primary, and secondary outcomes. *SRD*, Standard-risk donor.

Section 2

Questions and Answers about the Project

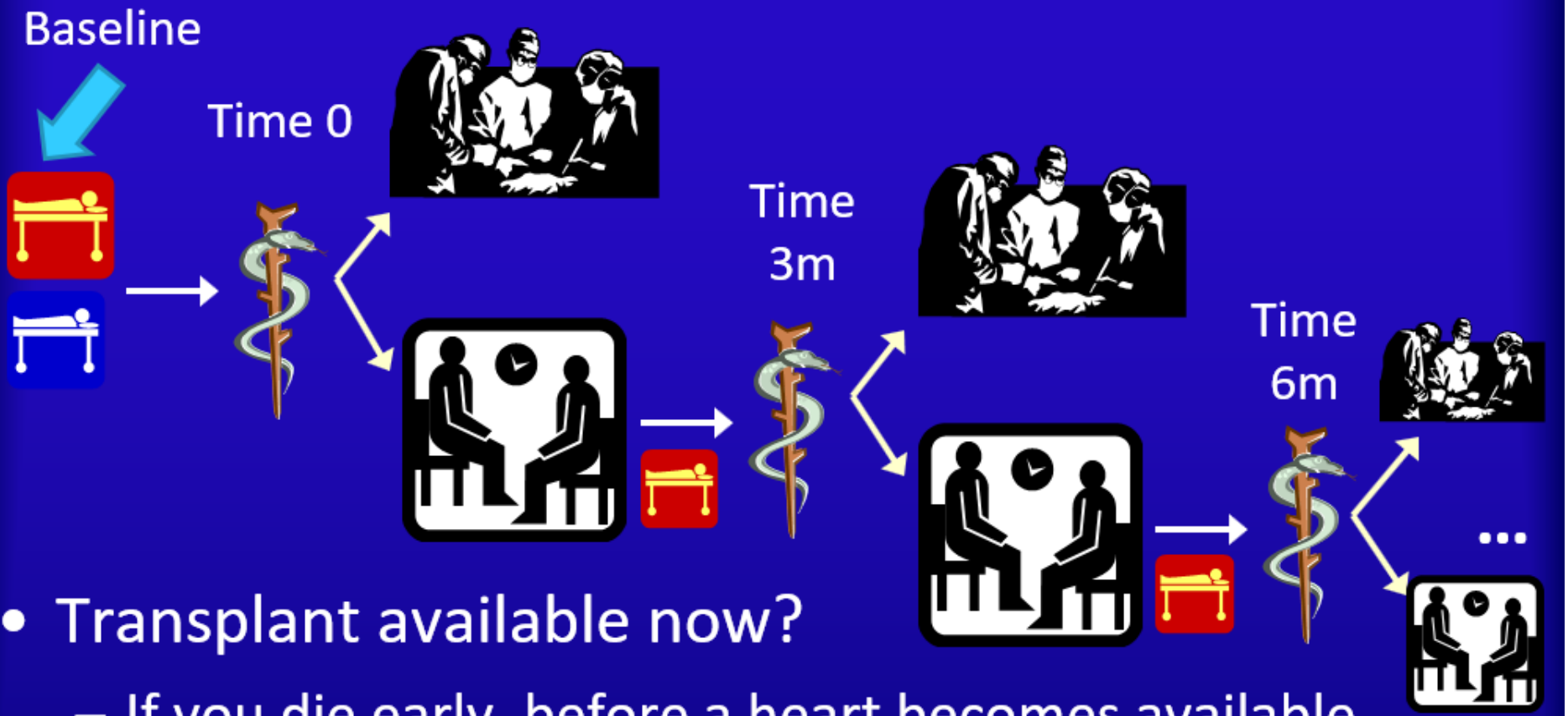
Section 3

Dealing with Time-Varying Covariates

When Do Time-Varying Covariates Arise?

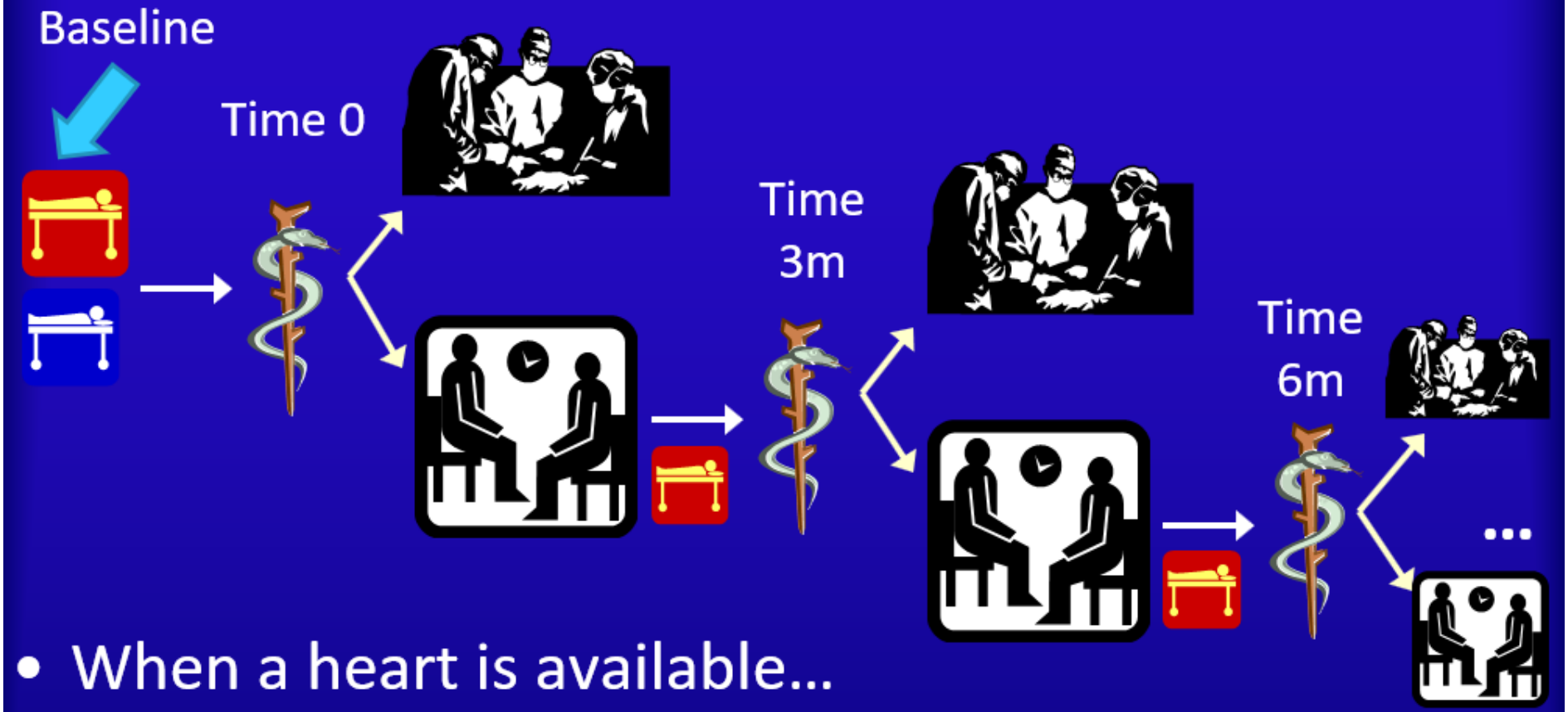
- When a treatment may be given at various times, we need to form matched pairs or sets in which subjects are similar prior to treatment, but avoid matching on events that followed treatment.
- In **risk-set matching**, a newly treated subject at time T is matched to controls who are not treated at time T based on covariate information describing subjects prior to time T.

Surgery Now vs. Waiting...



- Transplant available now?
  - If you die early, before a heart becomes available, you become a “control.”
  - If you live long enough to get a new heart, you’re treated.

Building Randomized Experiments



- When a heart is available…
  - Identify a pair of well-matched patients at that time
  - Randomly give heart to one, and compare survival…
  - Or keep “non-receiver” on list; investigate delay effects

Risk-Set Matching in a Study of Surgery for Interstitial Cystitis

- Effects of surgery (cyctoscopy & hydrodistention) on symptoms of IC (chronic urologic disorder)
  - IC Database - patients can enter after at least six months of IC symptoms
  - Patients evaluated at entry, then every 3 months
  - Pain (0-9), Urgency (0-9), Nocturnal Voiding Freq.
- Patients were treated periodically, not selected at random - those who find current symptoms unbearable will presumably opt for surgery

See Li, Propert and Rosenbaum 2001, Rosenbaum 2010, Chapter 12

Matching in the IC study

Can't compare all "surgery" to all "no surgery" patients since if you never had surgery, your symptoms probably were never intolerable.

- Want to create pairs of patients who were similar up to the moment one had surgery
- Matching makes pairs comparable pre-treatment
- What happens after treatment is an outcome
- Each new surgical patient is paired with a control who had similar symptoms up to the point of surgery for the surgical patient

What is the Exposure here?

We're estimating the effect of the choice that patients and their surgeons keep facing...

- Surgery Now vs. Delaying Surgery
  - with the delay being into the indefinite future

We're comparing Matched Pairs to estimate the impact of this decision

Match on Observed History

- Patient A has surgery at time T, with covariate values a

Find a matched control by selecting from those potential controls who:

- Have covariate history = a (or close to it)
- Who did not receive the treatment at time T

At time T, select a match for a patient getting surgery (with history a) from among the set of unmatched controls with similar history.

Matching Algorithm

- Use a distance metric as we've seen previously, but with a change.
- Distance between patient with surgery at time T and a patient who had not yet received surgery as of time T computed using covariates for those two patients up to time T
- Used no information obtained after time T.

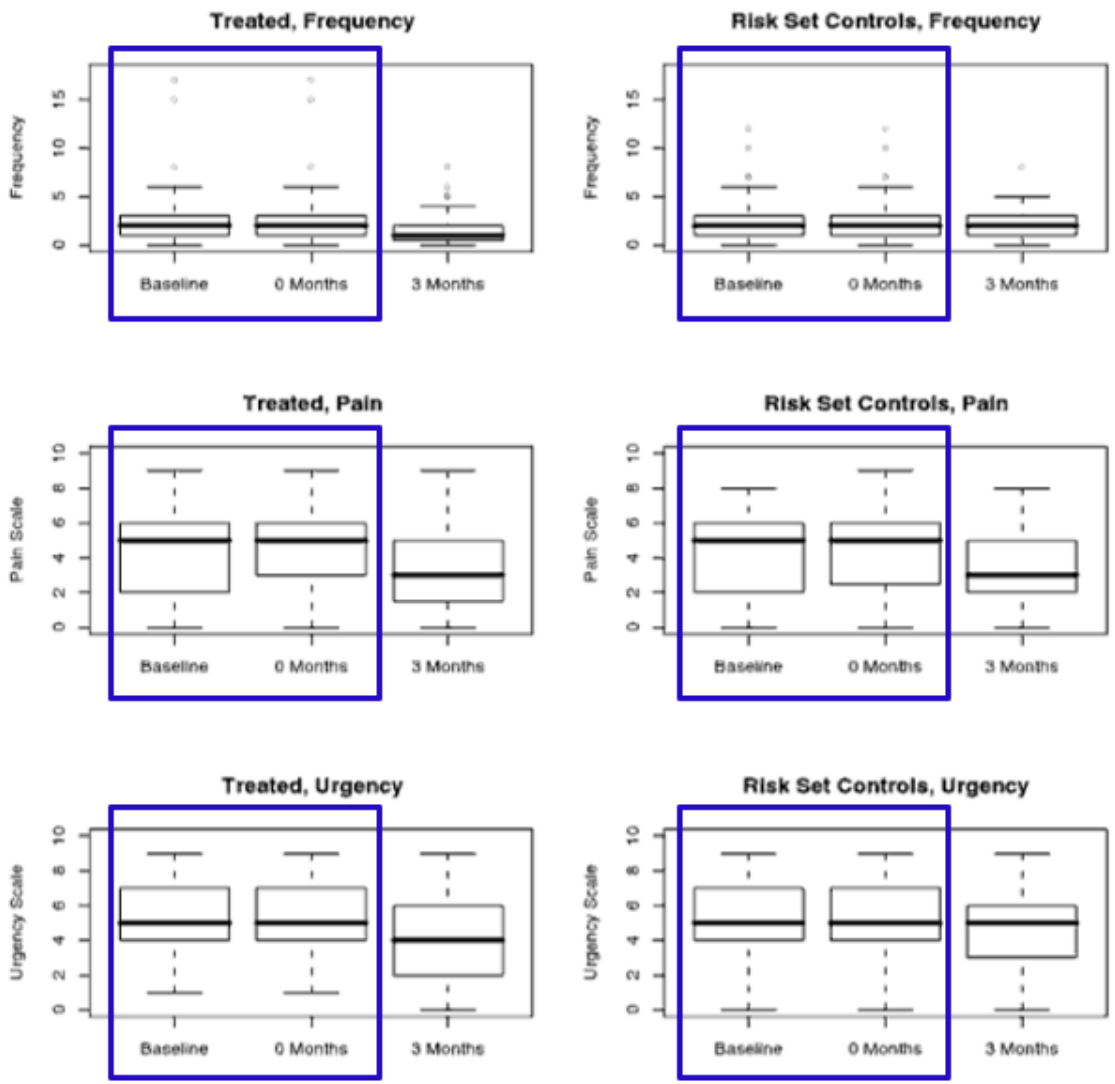


Fig. 12.1 Frequency, urgency, and pain at baseline, at treatment, and three months after treatment for treated patients and their matched not-yet-treated risk-set controls.

Effects of Matching

- Baseline
- 0 months (when treated)
- 3 months later
- Rank Corr's. > .9 at baseline and at 0 months btw treated & their matched controls

Rosenbaum 2010, p. 226

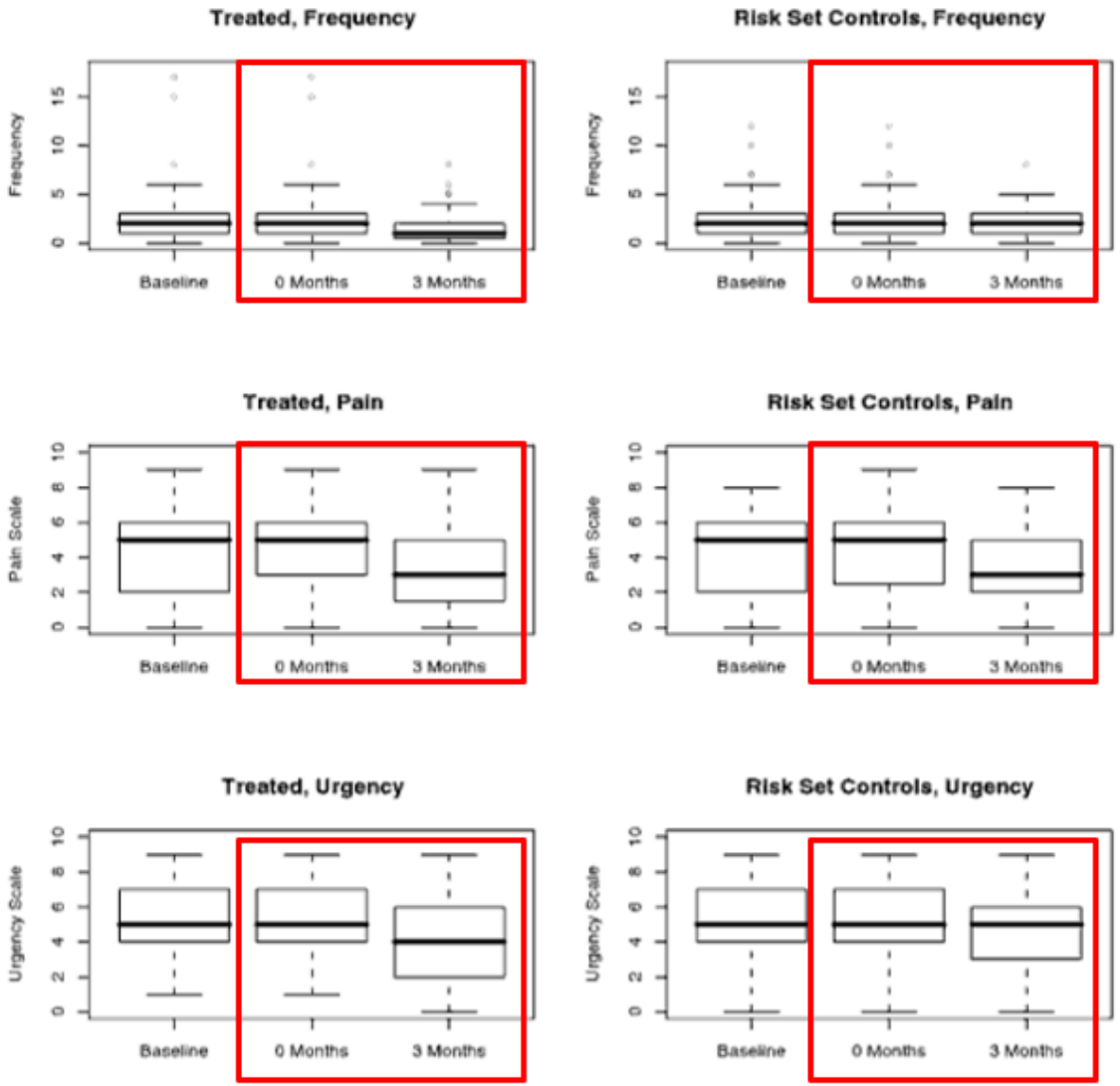


Fig. 12.1 Frequency, urgency, and pain at baseline, at treatment, and three months after treatment for treated patients and their matched not-yet-treated risk-set controls.

IC Conclusions

- Improved pain and urgency scores at 3 m for both treated and risk set controls.
- Frequency improved a bit for treated, not controls…

Rosenbaum 2010, p. 226

Another Example: Maturity at Discharge from Neonatal ICU

- Babies are kept in NICU until they have matured sufficiently to go home.
- Once babies look mature enough they stay around for "a few days" just to be sure.
- Does a longer stay in the NICU benefit those babies who receive it?

Silber et al 2009, Rosenbaum 2010 Section 12.3

NICU Maturity at Discharge Study

- 1402 premature infants (gestational age no more than 34 weeks) born at KP in Northern CA: 1998-2002
- Approach: risk-set matching using time-dependent propensity scores and an optimal nonbipartite matching
- Result: 1402 babies divided into 701 pairs so that one was an "early baby" and the other was a "late baby" but were similar at the postmenstrual age (to the day) that the "early baby" went home.

Available Covariate Information (1402 babies)

- 20 fixed (not time-dependent) covariates
  - Gestational age at birth, weight at birth, sex
  - Baby's health history
  - Mom covariates
- 13 time-dependent (maturity) covariates
  - Postmenstrual age, Current weight
  - Propensity to discharge
  - Six maturity measures scored daily (1 if not achieved, 0 if achieved) with exponential smoothing applied to the binary variables.

Time-Dependent Propensity Score

- Fit a Cox proportional hazards model to describe the hazard of discharge on a particular day, using as predictors both the fixed and time-dependent variables
- The linear portion (log-hazard of discharge) is then taken as the time-dependent PS
- The PS will thus vary from day to day for each baby.

Lu (2005) descibed the general approach.

Optimal Non-Bipartite Matching

- Non-bipartite matching doesn't presuppose treated patients and controls
- Begin instead with a single group and form pairs so as to minimize the distance between paired subjects or sets of matched subjects
- Very flexible - permits looking at doses, for instance, or matching with several groups

Lu, Zanutto et al. 2002, Rosenbaum 2010, Chapter 11

Distance Matrix for Nonbipartite Matching

Table 11.1 A 6×6 distance matrix for nonbipartite matching for six units. Unlike treatment-control matching, every unit appears as both a row and a column of this distance matrix. The optimal nonbipartite matching (1,2), (3,6), (4,5) is shown in **bold** with a minimum total distance of 106+25+34 = 165.

ID	1	2	3	4	5	6
1	0	**106**	119	231	110	101
2	**106**	0	207	126	192	68
3	119	207	0	156	247	**25**
4	231	126	156	0	**34**	67
5	110	192	247	**34**	0	212
6	101	68	**25**	67	212	0

- So, in the NICU study, we'll have 1402 rows and 1402 columns (1402 babies), and we'll be looking for the 701 pairs of babies with minimum total distance, under constraints.

Rosenbaum 2010, Chapter 11

Doing Risk-Set, Optimal Bipartite Matching in this Study

- Proportional hazards model gives time-dependent PS for each baby
- Built a 1402 x 1402 distance matrix comparing babies to one another.
- If the baby in row i and in column j were discharged on same day, used infinite distance.
- Otherwise, value in cell i, j describes distance between babies on the day earlier baby was discharged, with a caliper on the PS.

Table 12.2 Balance on fixed and time-dependent covariates after risk-set matching for 1402 premature babies in 701 matched pairs. Matching ensured that paired babies were similar on the day the early baby was discharged home from the neonatal intensive care unit, but the late baby was more mature (older, heavier) on the later day of discharge for the late baby. Of course, the fixed covariates are the same on both days; only the time-dependent covariates change.

Covariate Group	Covariate	Early Baby at Early Baby Discharge	Late Baby at Early Baby Discharge	Late Baby at Late Baby Discharge
	Number of Babies	701	701	701
Baby at Birth (fixed)	Gestational Age (weeks) at birth	31.1	31.1	31.1
	Weight at birth (grams)	1669	1686	1686
	SNAP-II 20 to 59	0.15	0.13	0.13
	SNAP-II 10 to 19	0.18	0.20	0.20
	SNAP-II 0 to 9	0.67	0.67	0.67
	Male Sex	0.51	0.52	0.52
Baby's Health History (fixed)	Bronchopulmonary Dysplasia	0.09	0.11	0.11
	Necrotizing Enterocolitis	0.01	0.01	0.01
	Retinopathy Stage ≥ 2	0.06	0.06	0.06
	Intraventricular Hemorrhage ≥ 3	0.02	0.01	0.01
Mom (fixed)	Maternal Age (years)	29.9	30.3	30.3
	Marital Status Single	0.24	0.24	0.24
	Other children = 0	0.40	0.37	0.37
	Other children = 1	0.34	0.37	0.37
	Other children ≥ 2	0.26	0.26	0.26
	Income $	59,517	59,460	59,460
	White Race	0.47	0.48	0.48
	Black	0.10	0.09	0.09
	Asian	0.20	0.23	0.23
	Hispanic	0.22	0.18	0.18
Baby's Time Dependent Variables	Postmenstrual Age (days)	247.4	247.4	250.9
	Propensity to discharge	0.67	0.64	1.33
	Apnea smoothed score	0.04	0.05	0.03
	Brady smoothed score	0.06	0.07	0.04
	Methyl smoothed score	0.04	0.03	0.02
	Oxygen smoothed score	0.11	0.11	0.07
	Gavage smoothed score	0.22	0.23	0.10
	Incubator smoothed score	0.15	0.15	0.08
	Combined maturity score	0.62	0.63	0.34
	Current weight	2153	2148	2231
	Current weight $<$ 1700 grams	0.02	0.03	0.01
	1700 $\leq$ weight $<$ 1800	0.06	0.06	0.02
	Current weight $\geq$ 1800 grams	0.92	0.91	0.97

Balance Achieved via Risk-Set Matching across the 701 pairs

Larger Version of Bottom Section on Next Slide

Rosenbaum 2010, p. 229

Balance Achieved on Time-Dependent Covariates

Covariate Group	Covariate	Early Baby at Early Baby Discharge	Late Baby at Early Baby Discharge	Late Baby at Late Baby Discharge
	Number of Babies	701	701	701
Baby's Time Dependent Variables	Postmenstrual Age (days)	247.4	247.4	250.9
	Propensity to discharge	0.67	0.64	1.33
	Apnea smoothed score	0.04	0.05	0.03
	Brady smoothed score	0.06	0.07	0.04
	Methyl smoothed score	0.04	0.03	0.02
	Oxygen smoothed score	0.11	0.11	0.07
	Gavage smoothed score	0.22	0.23	0.10
	Incubator smoothed score	0.15	0.15	0.08
	Combined maturity score	0.62	0.63	0.34
	Current weight	2153	2148	2231
	Current weight $<$ 1700 grams	0.02	0.03	0.01
	$1700 \leq$ weight $<$ 1800	0.06	0.06	0.02
	Current weight $\geq$ 1800 grams	0.92	0.91	0.97

Rosenbaum 2010, p. 229

Absolute Standardized Differences Across 20 Fixed & 13 Time-Dependent Covariates for 701 pairs

Quantile	Min	25	50	75	Max
Fixed	0.00	0.01	0.04	0.06	0.09
Time-Dependent (at early discharge)	0.00	0.01	0.02	0.06	0.09
Time-dependent (at own discharge)	0.09	0.16	0.19	0.34	0.75

- Babies are quite similar on the day the early baby was discharged – but quite different on the day of their own discharge

Silber et al. 2009, Rosenbaum 2010, Section 12.3

Conclusions?

Did the "late babies" benefit from a few more days to grow in the NICU?

- As it turns out, the "early" and "late" babies had similar experiences after discharge.
- No real strong indications of either benefit or harm for those with the extra time in NICU.
    - Not that NICUs aren't expensive places to be...

Silber et al. (2009)

Section 4

Augmenting a Sensivity Analysis (Rosenbaum 2017, Chapter 9)

Augmenting a Sensitivity Analysis

Lots of things can be described as part of a sensitivity analysis. We are focusing on one issue: quantifying departures from randomized (i.e. ignorable) treatment assignment.

Ignorable treatment assignment means that if two people have the same values of the observed covariates (and thus, for example, the same propensity score) then they have the same probability of treatment.

- As discussed in Chapter 9, Rosenbaum's bounds on Γ are just one possibility.
- Γ and Θ_p and Λ and Δ are just different methods of describing departure from ignorable treatment assignment in matched pairs, although only Γ applies outside of matched pairs.

Γ and Θ_p

We can express this in terms of Γ or Θ_p pretty easily in the matched pairs setting.

$$\frac{1}{1+\Gamma} \leq \Theta_p \leq \frac{\Gamma}{1+\Gamma}$$

$\Gamma = 2$ is the same magnitude of departure from ignorable treatment assignment as the interval from 0.33 to 0.67 for Θ_p.

If $\Gamma = 2$, then Harry might be twice as likely as Sally to receive the treatment (so Harry's probability Θ_H is $2/3$ and Sally's is $1/3$) or Sally might be twice as likely as Harry (so Harry's probability could be as low as $1/3$) to receive the treatment.

Amplifying the Γ value with Λ and Δ

This approach, like Θ_p bounds, applies only in the case of matched pairs.

- Λ tells you about the relationship of an unobserved covariate with treatment assignment.
- Δ tells you about the relationship of an unobserved covariate with the outcome.

$$\Gamma = (\Lambda\Delta + 1)/(\Lambda + \Delta)$$

Rosenbaum, 2017, Table 9.1

Table 9.1. Understanding the sensitivity parameter Γ

Γ	*Range of possible values of* θ_p		Λ	Δ
1	0.50	0.50	1	1
1.05	0.49	0.51	1.37	1.37
1.1	0.48	0.52	1.40	1.80
1.25	0.44	0.56	2	2
1.5	0.40	0.60	2	4
2	0.33	0.67	3	5
2.5	0.29	0.71	4	6
3	0.25	0.75	5	7
3.5	0.22	0.78	6	8
4	0.20	0.80	7	9
4.5	0.18	0.82	8	10
5	0.17	0.83	9	11
6	0.14	0.86	11	13
7	0.12	0.88	13	15
8	0.11	0.89	15	17
9	0.10	0.90	17	19
10	0.09	0.91	19	21

Using the Amplification

If $\Gamma = 1.5$ then, for example we could use

- a bound on Θ_p from 0.40 to 0.60
- or a combination of $\Lambda = 2$ and $\Delta = 4$
- or a combination of $\Lambda = 4$ and $\Delta = 2$
- or a requirement that $\Lambda = 1.5$ and that the unobserved covariate be a perfect predictor of the outcome.
- or a requirement that $\Delta = 1.5$ and that the unobserved covariate be a perfect predictor of treatment assignment.

Summary: Sensitivity Analysis

Hidden bias is the great problem with observational studies, and with PS models.

- Sensitivity analysis after matching can be applied in many scenarios.
- We hope to find that an unobserved covariate would have to be very powerful to alter our conclusions.
- That doesn't mean that such a covariate (or set of them) doesn't exist.
- You should be running a sensitivity analysis in your Project only if your matched sample results show an effect size which implies an impact for your exposure.