# Does transfer to intensive care units reduce mortality for deteriorating ward patients?

## Estimating Person-centered Treatment (PeT) Effects Using Instrumental Variables

Anirban Basu
***University of Washington, Seattle***

*(Jointwork with **Stephen O'Neill**, National University of Ireland Galway and **Richard Grieve**, London School of Hygiene and Tropical Medicine)..... Special thanks to **Steve Harris***

UNIVERSITY *of* WASHINGTON

# Treatment Evaluations

- Literature on treatments effects focused on informing population level or policy-level decisions.
  - Mean effects
  - Distributional effects
  - Impacts are viewed as informing a social decision maker to help choose across alternative options
  - Discussions will focus on evaluating policies/treatment in use

- Heterogeneity: Optimal Individual choices may vary from socially optimal choices, based on *ex-post* realizations.

# Literature on Distribution of Treatment Effects



- Imbens and Rubin (1997) and Abadie (2002, 2003)
- Carniero and Lee (2009)
- Heckman and Honoré 1990, Heckman and Smith 1993, Heckman et al. 1997.
- Factor structure models have been used to establish the joint distribution of potential outcomes (Aakvik et al. 1999, Carniero et al. 2003).
- Importance of distribution of effects well established in the literature (Abbring and Heckman, 2007)

# Healthcare Setting

- Dilemma between social versus individual choices, perhaps, the most stark
- In traditional clinical outcomes research, the focus has always been on finding average effects either through large clinical trials or observational datasets.
- Estimating treatment effect heterogeneity has mostly been relegated to post-hoc analysis, rather than becoming the central goal of the analysis.
- Growing recognition of the importance of nuanced and possibly individualized estimates of treatment effects (Basu 2009, 2011, Basu et al. 2011).

# Observational Data Setting

- Observational data is a valuable resource.
- Selection bias in observational studies is a potential problem – primary reason for relegation to second-tier status in evidentiary standards.
- However, growing interests due to large scale investments in Electronic Health Records.
- Methods for casual inference in such data of high demand.

# Outline

- Definitions, identification and estimation of treatment effects using instrumental variable methods
- PeT effects
- PeT effects estimation
- Empirical Example

# *Potential Outcomes Model*

*Structural models of outcomes and treatment choice following Heckman and Vytlacil (1999, 2001, 2004).*

*The treated state denoted by j = 1 and the untreated state denoted by j = 0,*

*Assume,*

$$Y_1 = \mu_1(X_O, X_U, \vartheta) \quad \textit{and} \quad Y_0 = \mu_0(X_O, X_U, \vartheta) \qquad (1)$$

***Assumption 1***. $(X_O, X_U) \amalg \vartheta$ *where* $\amalg$ *denotes statistical independence.*

# Treatment Choice Model

*Individual choose to be in state 1 or 0 (prior to the realization of the outcome of interest) according to the following equation:*

$$D = 1 \text{ if } \mu_D(X_O, Z) - U_D > 0 \qquad (2)$$

*By definition,* $U_D \perp\!\!\!\perp \vartheta$, *which also defines the distinction between* $X_U$ *and* $\vartheta$ *in (1).*

*Assumption 2:* $(X_U, U_D) \perp\!\!\!\perp Z \mid X_O$
*Assumption 3:* $\mu_D(x_O, Z)$ *is nondegenerate*
*Assumption 4:* $U_D$ *is continuous*
*Assumption 5:* $E(Y_1)$ *and* $E(Y_0) < \infty$
*Assumption 6:* $1 > Pr(D=1 \mid X_O) > 0$

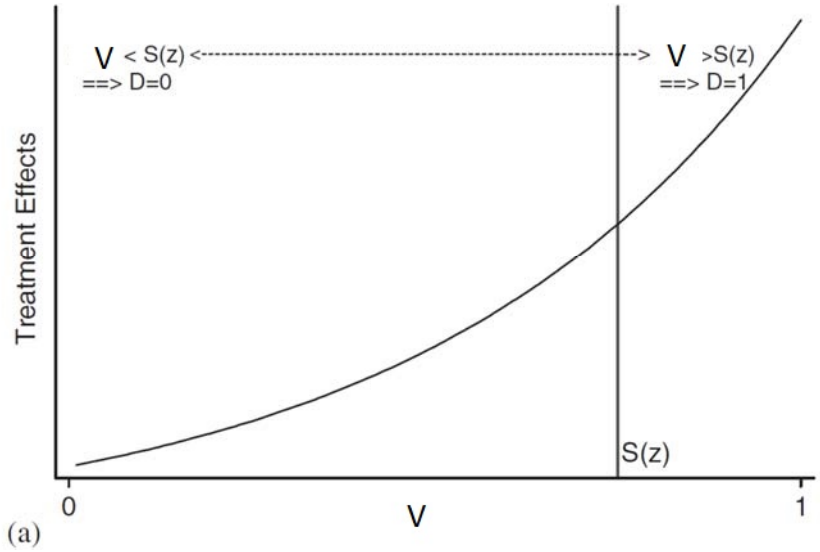
*Rewrite (2) as*

$$D = 1 \text{ if } P(X_O = x_O, Z = z) > V \qquad (3)$$

*where*

$$V = F_{U_D}[U_D | X_O = x_O, Z = z],$$

$$= F_{U_D}[\mu_D(x_O, z)].$$




*Therefore, for any arbitrary distribution of* $U_D$ *conditional on* $X_O$ *and* $Z$*, by definition,* $V \sim Unif[0, 1]$ *conditional on* $X_O$ *and* $Z$.

# *Definitions of Treatment Effects*

*An individual–level treatment effect is given as*

$$TE = (Y_1 - Y_0) \qquad (4)$$

*TE differs across individual subjects depending on* $X_O$, $X_U$, *and* $\vartheta$.

$\vartheta$ *is typically not only unmeasured but also unknown (as otherwise would have been used for treatment selection),*

# *Definitions of Treatment Effects*

*The most accurate individualized expected treatment effect (IETE):*

$$\xi(x_O, x_U) = E_\vartheta(Y_1 - Y_0 | x_O, x_U) \qquad (5)$$

*Typically, since only* $X_O$ *are observed, a conditional average treatment effect (CATE) (Heckman 1997) can be formed:*

$$CATE = \xi(x_O) = E_{X_U} E_\vartheta(Y_1 - Y_0 \mid x_0), \qquad (6)$$

# *Estimation*

$$Y = Y_1 \quad \textit{if } D = 1 \textit{ and } Y = Y_0 \quad \textit{if } D = 0 \qquad (8)$$

*Goal of the analysis – estimate*

$Y_1$ *for subjects with* $D = 0$

$Y_0$ *for subjects with* $D = 1$.

*Overt Selection Bias:*

- *Differences in* $X_O$ *across treatment groups.*
- *Address with statistical methods such as covariate adjustment, propensity score matching, etc*

*Hidden Selection Bias:*

- *Differences in* $X_U$ *across treatment groups.*
- *Formally address with Selection models or IV methods.*

# *Instrumental Variables (IV)*

*Instrumental variables determine treatment choice, but do not affect outcomes other than through their effects on treatment choice.*

*IVs can induce substantial variation in the treatment variable but have no direct effect on the outcome variable of interest.*

*Estimate how much the variation in the treatment variable that is induced by the instrument—and only that induced variation—affects the outcome measure.*

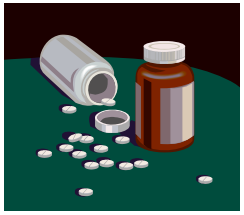
*Leads to the traditional IV estimator.*

Unobserved confounder:
Fundamental problem of evaluation

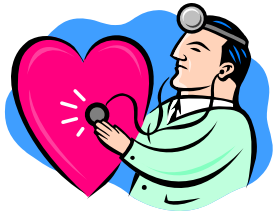
Confounder

IV methods

Instrumental variable



Exposure/ Treatment



Outcomes

IV methods

Confounder

# *Assumptions on Heterogeneity for Traditional IV*

*Key Assumption:*

*Effects of treatment do not vary by levels of* $X_U$.

*IV estimator then estimates the Average Treatment Effect for a population if:*

*a) treatment effects are constant for every one in the population with the same observed characteristics or*

*b) even if treatment effects are heterogeneous, individuals do not have any additional information beyond what the analyst of an observational data possess that can enable them to anticipate these effects and select into treatment that would potentially give them the largest benefits.*

# *Selection Bias and Heterogeneity*

*Relaxing these assumptions imply:*

*Treatment effects vary by levels of* $X_U$ *(unobserved heterogeneity) & this heterogeneity correlated with treatment selection, <u>because</u>*

*Subjects choose treatment anticipating idiosyncratic gains that are not observed by analyst (Heckman 1997)*

*OR subjects cannot anticipate idiosyncratic gains but select treatment based on* $X_U$*, which determines treatment heterogeneity (Basu 2011).*

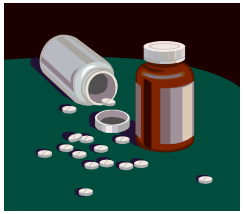
➔ essential heterogeneity *(Heckman 1997; Heckman et. al., 2006)*

Unobserved confounder:
Fundamental problem of evaluation

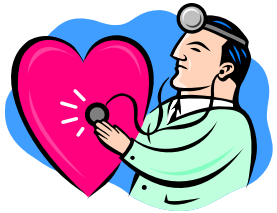
IV methods

Confounder

Instrumental variable



Exposure/ Treatment



Outcomes

IV methods

Confounder

# *Selection Bias and Heterogeneity*

*Imbens and Angrist (1994) showed that when the standard assumptions are relaxed, standard IV methods can identify parameters that reflect the treatment effects for a* specific *group of patients:*

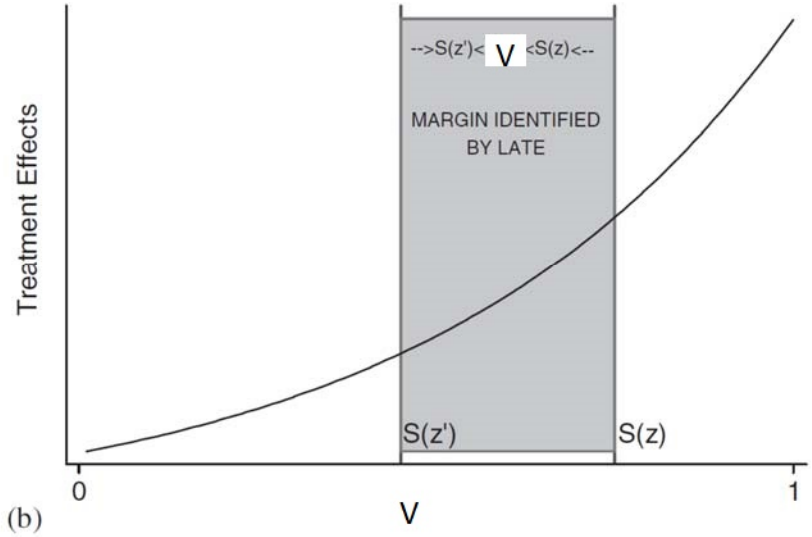
> *who would change treatment in response to changes in the levels of the instrument*

*Only under a choice-theoretic formulation, one can attribute a marginal interpretation to these patients.*

# *Definitions of Treatment Effects*

Local Average Treatment Effect (LATE)

$$E\,(Y_1 - Y_0 \mid D(z) = 1,\ D(z_0) = 0)$$
$$= E\,(Y_1 - Y_0 \mid \upsilon_0 < V < \upsilon)$$
$$= LATE(\upsilon, \upsilon_0)$$




$$\upsilon = Pr(D\,(z) = 1) = Pr(D = 1 \mid Z = z) = P\,(z),$$
$$\upsilon_0 = Pr(D\,(z0) = 1 \mid Z = z0) = Pr(D\,(z0) = 1) = P\,(z0).$$

# *Selection Bias and Heterogeneity*

*The issue then is that the marginal patients identified by the IV methods are entirely dependent on the specific instrument being used and how this instrument affects treatment choices:*

- ➔ *the use of different instruments will produce different treatment effects because they represent the effects for different groups of marginal patients*
- ➔ *IV results become instrument dependent*
- ➔ *Hausman test of over-identification no longer applies*

*More importantly, these marginal patients are not readily identifiable!! (McClellan & Newhouse 1998, Harris & Remler 1998, Brooks et al. 2004)*

# *Definitions of Treatment Effects*

Marginal Treatment Effects (MTE)

*The MTE is the average gain to patients who are indifferent between receiving treatment 1 versus treatment 0 given* $X_O$ *and Z.*

$$\text{MTE}(X_O = x_O, V = v) = \boldsymbol{E_{X_U|P(x_O,z)=v} E_{\vartheta}(Y_1 - Y_0 | x_O, x_U)}$$
$$= \boldsymbol{E_{\vartheta}(Y_1 - Y_0 | x_O, v)|_{P(x_O,z)=v}} \qquad (7)$$

*MTE are nuanced parameters, can be aggregated to form any of the mean treatment effect parameter, but lacks individual identity.*

# LIV Estimator

1st stage

$$\text{logit}(D) = \alpha_0 + \alpha_1 X + \alpha_2 Z$$

2nd stage

$$E(Y \mid X, \hat{p}(x,z)) = g(\beta_0 + \beta_1 X + \beta_2 x \cdot \hat{p} + K(\hat{p}; \underset{\sim}{\beta_3}))$$

$$MTE(x, U_D = u_D) = \left. \frac{d\hat{E}(Y \mid X, \hat{P}(x,z))}{d\hat{P}} \right|_{\hat{P}=p=(1-u_D)}$$

$U_D$ = Propensity for treatment selection based on unobserved confounders

# Extension to Person-centered Treatment (PeT) effect



Marginal patients:

$$D^* = g(\text{Distance}) + h(\text{Severity}) = 0$$

Therefore, the severity level of a marginal patients:

$$\text{Severity} = h^{-1}(-g(\text{Distance}))$$

Standard calculus can allow one to characterize the level of severity as a function of the instrument level, where the marginal patient is identified.

Why do we care?

# Extension to Person-centered Treatment (PeT) effect



Why do we care?

Go back to data.

Take patient 1

- Observe Z, D
- Calculate the range of Severities for which D satisfies D* = g(Z) + h(Severity)
- Average the MTE over only those range of severities for patient 1

Obtain a person-centered treatment effect for patient 1!

# *Definitions of Treatment Effects*

Person-centered Treatment (PeT) Effects

*Person-centered Treatment (PeT) effects, can be written as:*

$$\Delta = \boldsymbol{E}_{\boldsymbol{X_U}|\boldsymbol{P}(\boldsymbol{x_O},Z),D} E_{\vartheta}(Y_1 - Y_0 | x_O, x_U) \qquad (8)$$

*where the expectation of unobserved confounders is made conditional on person-specific estimates of* $X_O$, $P(Z)$ *and* D.

*e.g. for D = 1:*

$$\Delta(1) = \boldsymbol{E}_{\boldsymbol{X_U}|V<\boldsymbol{P}(\boldsymbol{x_O},Z)} E_{\vartheta}(Y_1 - Y_0 | x_O, x_U) = P(z)^{-1} \int_0^{P(z)} MTE(x_O, v) dv$$

# *Uses of PeT effects*

$$ATE = E_{X_O,P(Z),D}(\Delta)$$
$$TT = E_{X_O,P(Z)|D=1}(\Delta) \qquad (7)$$
$$TUT = E_{X_O,P(Z)|D=0}(\Delta)$$

*Can also be used to forecast policy effects for any policy that shifts a certain subgroup of individuals, characterized by shifting the distribution of* $X_O$*, to take up or give up treatment.*

# *Uses of PeT effects*

- *They help to comprehend individual-level treatment effect heterogeneity better than CATEs.*

- *They are better indicators for the degree of self-selection than CATE. Specifically,*

  - *Better predictors of $\xi(x_O, x_U)$ both in terms of the positive predictive value $Pr(\xi(x_O, x_U) \geq 0 \mid \Delta \geq 0)$ and the negative predictive value $(Pr(\xi(x_O, x_U) < 0 \mid \Delta < 0))$, compared to CATEs.*

- *They can explain a larger fraction of the individual-level variability in treatment effects than the CATEs.*

# Estimation of PeT effects

- Basu A. Person-Centered Treatment (PeT) effects using instrumental variables: An application to evaluating prostate cancer treatments. ***Journal of Applied Econometrics*** 2014; 29:671-691.
- Basu A. Person-centered treatment (PeT) effects: Individualized treatment effects with instrumental variables. ***Stata Journal*** In Press.

- -petiv- command in STATA

# PeT Estimation Algorithm

- Numerical integration: For each individual *i*:
  - Draw 1000 deviates *u*~Uniform[*min(P), max(P)*]
  - Compute $d\hat{g}(.)/d\hat{p}$ and evaluate it by replacing P with each value of *u*.
  - Compute $D^* = \Phi^{-1}(P)+\Phi^{-1}(1-u)$ also generating 1000 values for each individual.
  - Compute PeT by averaging $d\hat{g}(.)/d\hat{p}$ over values of *u* for which ($D^* > 0$) if D=1, otherwise, by averaging over values of *u* for which ($D^* \leq 0$) if D=0.
- Estimated PeT effects provide us with individualized effects of treatment effects. Mean treatment effect parameters were also computed. Averaging PeTs over all observation gave ATE. Averaging PeTs over over D=1 or D=0 gave us TT and TUT
- respectively.

# Empirical Example

• • •

**Does transfer to intensive care units reduce mortality for deteriorating ward patients?**

# Background

- Adult Intensive Care Units (ICU) costly and scarce resource
  - Supply usually lags demand
- No RCT evidence
- Observational study evidence
  - Do not deal with the endogeneity of transfer
  - Do not recognizing heterogeneity in returns from transfer
- Transfers to ICU typically relies on clinical judgement
  - Not perfect proxy for reliable and causal evidence

# Our Study

- Exploit natural variation in ICU transfer according to ICU bed availability for deteriorating ward patients in the UK
- The (SPOT)light Study (N = 15,158)
  - Prospective cohort study of the deteriorating ward patients referred for assessment for ICU transfer
  - Hospitals were eligible for inclusion if they participated in the ICNARC Case Mix Programme
  - Patients recruited between Nov 1, 2010 - Dec 31, 2011 from 49 UK NHS hospitals
  - A variety of exclusion conditions were applied to identify deteriorating ward patients who are equipoised to be transferred to ICU

# Data

- Database locked on Sep 2012
- Detailed demographic and physiological data (collected from the time of ward assessment)
- Date of death (NHS Information Service)
- Critical care provision, ICU bed availability, and hospital characteristics (CMP and Hospital Episode Statistics (HES))
- Number of available ICU beds =

(Maximum number of beds reported to ICNARC) – (Number of actively treated patients occupying those beds at the time of ward assessment)

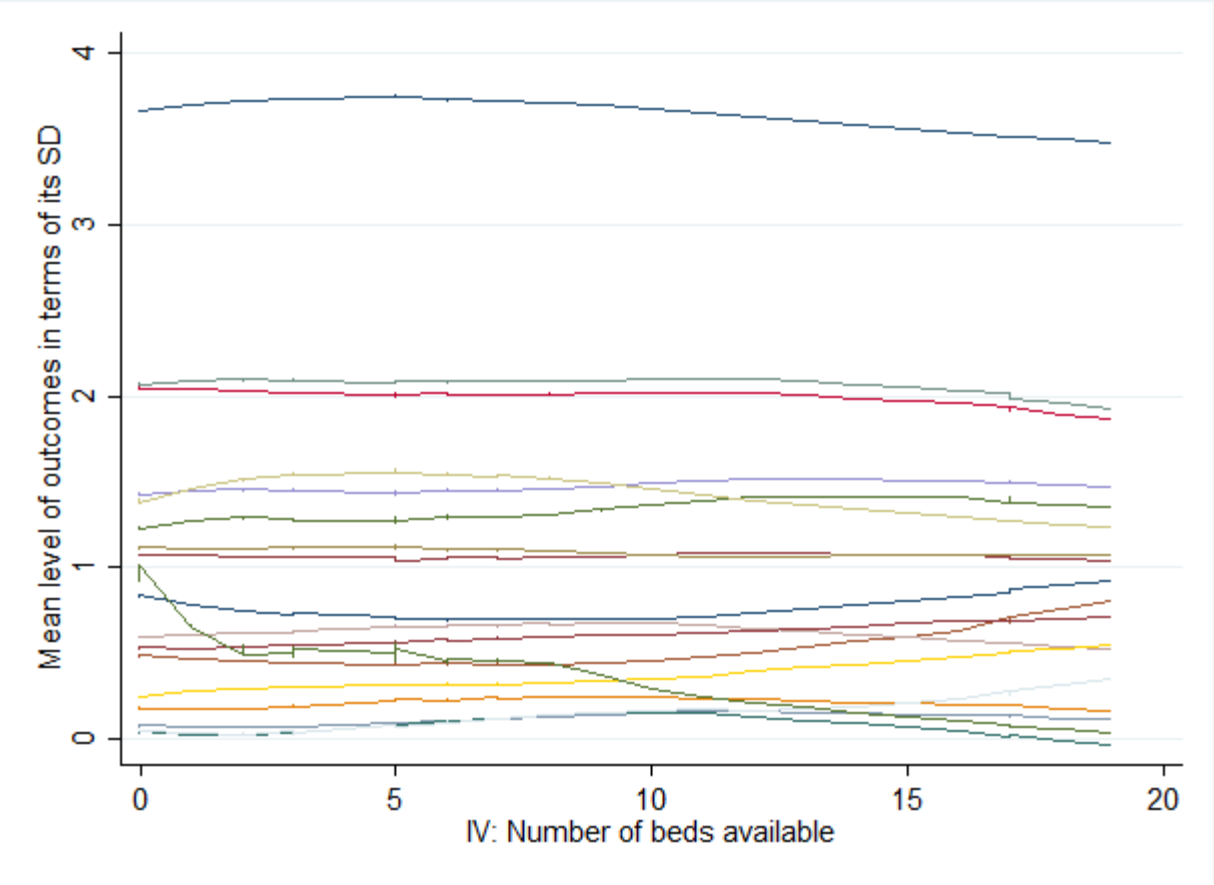
# Data

- Primary Outcome: Death 7 days post assessment
- Secondary Outcomes: Death within 28 and 90 days
- Exposure: ICU transfer vs care on general wards
- Baseline covariates: Age, diagnosis of sepsis, peri-arrest, dependency at ward assessment and recommended level of care post assessment (4 levels) and three physiology measures
  - National Early Warning Score (NEWS) : whether respiratory rate, oxygen saturations, temperature, systolic blood pressure, pulse rate, a level of consciousness vary from the norm,
  - the Sequential Organ Failure Assessment (SOFA), and
  - the ICNARC physiology score

# IV

- IV = NBA: Vary across hospital and over time
- Key Assumptions:
  - NBA at ward patient's assessment directly affects one's probability of transfer to ICU
  - NBA unconditionally independent of mortality of patients

# IV balance in unmatched

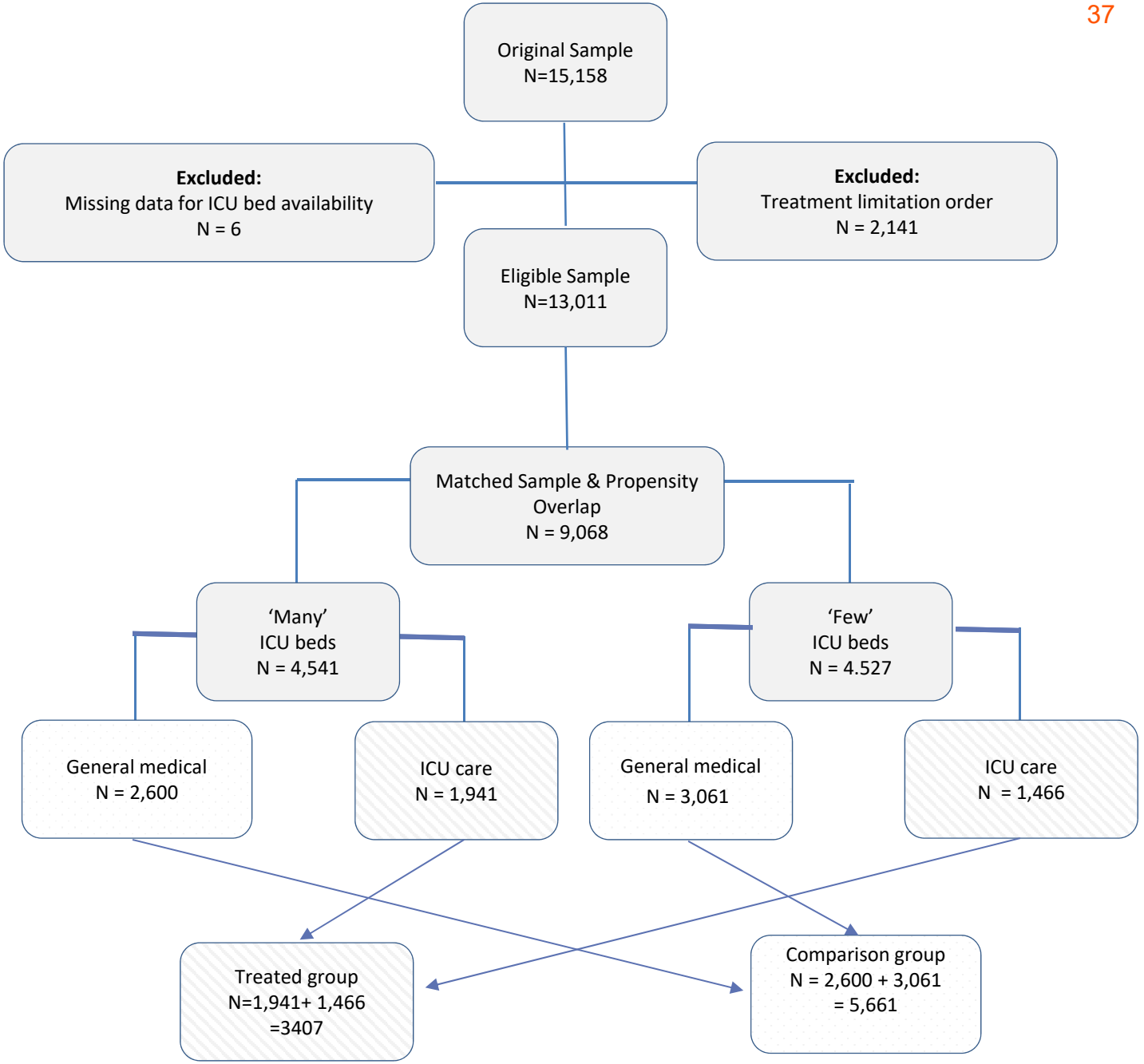


# IV and Near-Far Matching

- IV = NBA: Vary across hospital and over time
- Key Assumptions:
  - NBA at ward patient's assessment directly affects one's probability of transfer to ICU
  - NBA unconditionally independent of mortality of patients
- Increase strength of instrument – Near –far matching
  - Match similar patients with 'many' versus 'few' NBA
  - Similarity assessed on: age, gender, NEWS SOFA and ICNARC physiology scores, CCMDS level at assessment, and timing (out of hours, winter, and weekend or not)
  - Keep matched pairs with at least difference of 3 of more NBA

Baiocchi et al. JASA 2010; Small and Rosenbaum JASA 2008; Keele and Morgan AAS 2016

# Sample Selection Summary



# Near-Far Matching Results



| | BEFORE MATCHING | | | | | AFTER MATCHING |
|---|---|---|---|---|---|---|
| | ICU | | General medical | | Std. Diff | Std. Diff |
| No. of admissions | 4,994 | | 8,017 | | | |
| No. of ICU beds available | | | | | | |
| Mean (SD) | 4.63 | (3.22) | 4.05 | (3.13) | 0.181 | 2.905 |
| Median (Min, Max) | 4 | (0, 18) | 3 | (0, 19) | | |
| Age, mean (SD) | 63.77 | (16.74) | 66.08 | (18.30) | -0.132 | 0.013 |
| Male sex, n (%) | 2,728 | (54.6%) | 4,105 | (51.2%) | 0.069 | -0.010 |
| Reported sepsis diagnosis, n (%) | 3,380 | (67.7%) | 4,553 | (56.8%) | 0.226 | -0.002 |
| CCMDS level of care at visit, n (%) | | | | | | |
| Level 0 | 490 | (9.8%) | 1,231 | (15.4%) | -0.168 | -0.072 |
| Level 1 | 3,064 | (61.4%) | 5,774 | (72.0%) | -0.228 | 0.040 |
| Level 2 | 1,301 | (26.1%) | 944 | (11.8%) | 0.371 | -0.022 |
| Level 3 | 104 | (2.1%) | 22 | (0.3%) | 0.168 | 0.055 |
| Missing | 35 | (0.7%) | 46 | (0.6%) | 0.016 | 0.105 |
| Recommended CCMDS level of care at visit, n (%) | | | | | | |
| Level 0 | 86 | (1.7%) | 838 | (10.5%) | -0.371 | 0.174 |
| Level 1 | 1,183 | (23.7%) | 5,830 | (72.7%) | -1.126 | -0.004 |
| Level 2 | 2,539 | (50.8%) | 1,229 | (15.3%) | 0.815 | -0.046 |
| Level 3 | 1,152 | (23.1%) | 52 | (0.6%) | 0.739 | -0.036 |
| Missing | 34 | (0.7%) | 68 | (0.8%) | -0.019 | -0.106 |
| Peri-arrest, n (%) | 456 | (9.1%) | 191 | (2.4%) | 0.293 | 0.074 |
| Acute Physiology scores, mean (SD) | | | | | | |
| ICNARC | 17.40 | (7.76) | 13.63 | (6.55) | 0.524 | -0.021 |
| SOFA | 3.90 | (2.33) | 2.68 | (1.95) | 0.565 | -0.011 |
| NEWS | 7.07 | (3.17) | 5.68 | (2.93) | 0.456 | -0.030 |
| NEWS Risk class, n (%) | | | | | | |
| None | 93 | (1.9%) | 258 | (3.2%) | -0.086 | -0.015 |
| Low | 956 | (19.1%) | 2,451 | (30.6%) | -0.267 | 0.047 |
| Medium | 1,240 | (24.8%) | 2,487 | (31.0%) | -0.138 | -0.008 |
| High | 2,705 | (54.2%) | 2,821 | (35.2%) | 0.389 | -0.029 |
| Time of admission, n (%) | | | | | | |
| Weekend | 1,261 | (25.3%) | 1,969 | (24.6%) | 0.016 | 0.057 |
| Out of hours | 1,983 | (39.7%) | 2,608 | (32.5%) | 0.150 | -0.046 |
| Winter | 1,299 | (26.0%) | 2,064 | (25.7%) | 0.006 | -0.000 |

FAR

NEAR

# References

Basu A, Meltzer D. Value of information on preference heterogeneity and individualized care. Medical Decision Making 2007; 27(2):112-127.

Basu A, Heckman J, Navarro-Lozano S, et al. Use of instrumental variables in the presence of heterogeneity and self-selection: An application to treatments of breast cancer patients. Health Econ 2007; 16(11): 1133 -1157.

Basu A. 2009. Individualization at the heart of comparative effectiveness research: The time for i-CER has come. Medical Decision Making, 29(6): N9-N11.

Basu A, Jena AB, Philipson TJ. The impact of comparative effectiveness research on health and health care spending. J Health Econ. 2011;30(4):695-706.

Basu A. Economics of individualization in comparative effectiveness research and a basis for a patient-centered health care. J Health Econ. 2011; 30(3):549-559.

Basu A. Person-Centered Treatment (PeT) effects using instrumental variables: An application to evaluating prostate cancer treatments. Journal of Applied Econometrics (In Press).

Heckman JJ, Vytlacil EJ. Local instrumental variables and latent variable models for identifying and bounding treatment effects. Proc Nat Acad Sci 1999; 96(8): 4730-34

Heckman JJ, Urzua S, Vytlacil E. Understanding instrumental variables in models with essential heterogeneity. Rev Econ Stat 2006; 88(3): 389-432.

Kaplan S, Billimek J, Sorkin D, Ngo-Metzger Q, Greenfield S. Who Can Respond to Treatment?: Identifying Patient Characteristics Related to Heterogeneity of Treatment Effects. Medical Care 2010; 48(6): S9-S16