Confounding by indication developments in matching, and instrumental variable methods

Richard Grieve
London School of Hygiene and Tropical Medicine

Outline

1. Causal inference and confounding
2. Genetic Matching
3. Developments in Instrumental variable approaches
4. Conclusions and further research

Key Concern: confounding

HRT triples the risk of breast cancer, biggest ever study shows

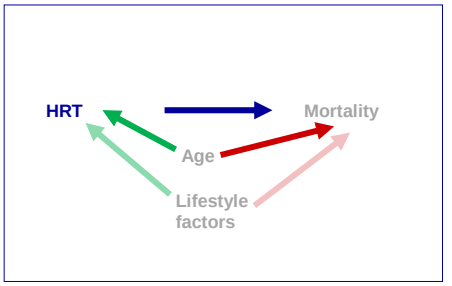
The Telegraph

'Hormone replacement therapy can triple the risk of breast cancer, the biggest ever study has found, following more than a decade of controversy. Last year NICE changed guidance to encourage more doctors to prescribe HRT claiming too many menopausal women had been left suffering in silence. Doctors were reluctant to prescribe it after a study in 2002 suggested it could raise the risk of cancer, a claim later widely disputed.
Now new findings suggest the original risk had actually been underestimated.'

A study of 100,000 women over 40 years found those who took the combined oestrogen and progestogen pill for around five years were 2.7 times more likely to develop cancer compared to women who took nothing, or only the oestrogen pill.

August, 2016

Selection bias from measured and unmeasured confounders

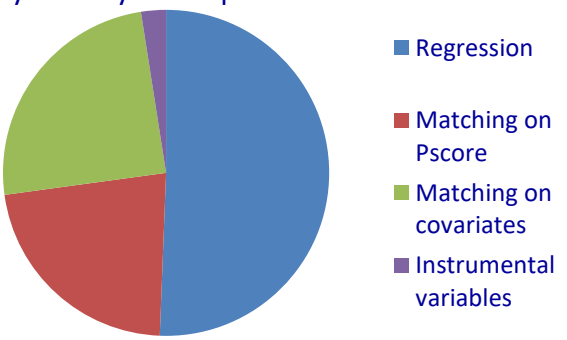


If ignored can lead to selection bias

- Bias from imbalance on unobservables: hidden bias
- Bias from imbalance on observables: overt bias

CEA that use observational data

- Systematic review econ eval (2000-2011) by Kreif et al (2012)
- 79 studies
- Most studies assumed "no unobserved confounding"
- Failed to justify this key assumption



Problem of causal inference (Rubin 1977, Holland 1986)

- T is treatment indicator: 1 treatment group, 0 control
- Interested in causal relationship between *T* and *Y*
- Potential outcomes *Y(0)* and *Y(1)* under control and treated states
- Ideally observe treatment effect for each individual $\tau_i = Y_i(1) - Y_i(0)$
- BUT cannot observe both outcomes
- **Objective of methods: impute missing potential outcome**

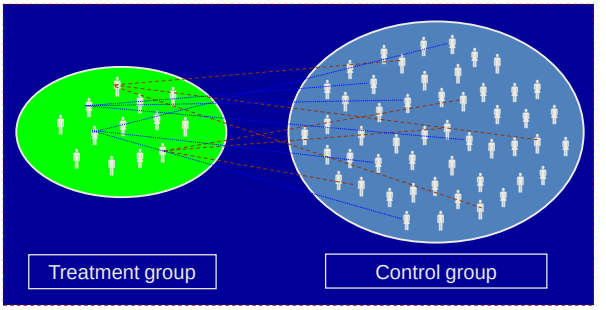
Statistical Methods for addressing confounding

- **Assume no unobserved confounding**
  - Regression adjustment
  - Matching and propensity score methods

  Assumptions: - no unobserved confounding
  - overlap (positivity)
- **Allow for observed and unobserved confounding:**
  - Instrumental variable estimation
  - Regression discontinuity design

Intuition behind matching

e.g. for estimating average effect of treatment on the treated



Require matching method that achieves **best balance** in observed characteristics x_i between treatment and control groups

Propensity score (PS) methods

- Most non-parametric way match exactly on X
- Feasible if very few, discrete confounders

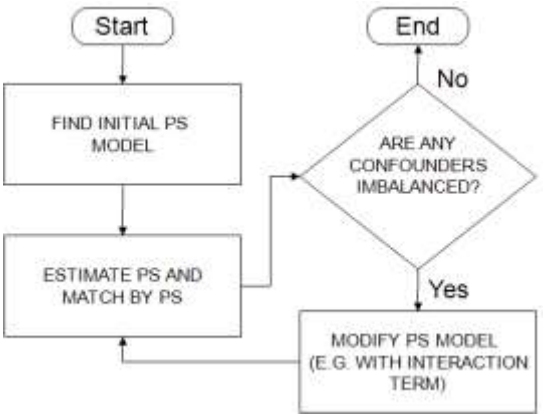
- Alternative: Propensity score methods (Rosenbaum and Rubin, 1983)

$$p(X_i) = \Pr(T_i = 1 \mid X_i)$$

- Model of the probability of treatment, given covariates
- Choice of treatment depends on patient, clinician choice
- Key result: PS is a balancing score
  - Sufficient to control for PS only
  - Matching, inverse weighting, subclassification, adjustment
- Specification test is assessment covariate balance

9

Iterative process for specifying the Pscore



10

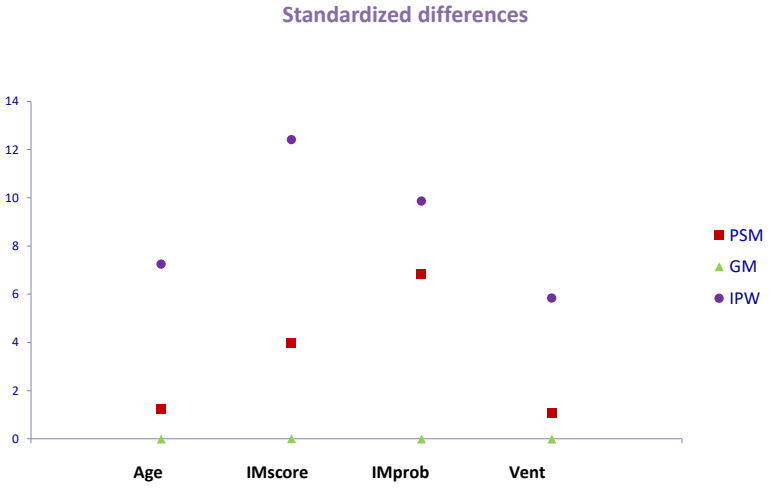
Genetic Matching (GM): Multivariate matching
(see Sekhon 2011, Sekhon and Grieve, 2011)

- GM extends traditional multivariate matching methods
- Uses general distance measure
- $GMD(X_i,X_j) = \{ (X_i - X_j)' (S^{-1/2})' \boldsymbol{W} S^{-1/2}(X_i - X_j) \}^{1/2}$
  - $\boldsymbol{W}$ is a weight matrix
- GM considers many alternative sets of weights
- A genetic algorithm searches data to pick the weights W
- In each iteration: creates matched pairs according to *GMD,* using chosen weights, checks balance
- Picks those weights that maximise overall covariate balance
- **Creates matched dataset using optimal weights**

Main features of Genetic Matching

- Specify variables want to match on (*X* matrix)
- Specify variables vital to balance (balance matrix)
- Choose balance statistics (e.g. t-tests, KS statistics)
- Specify matching options (e.g. 1 to 1, replacement)
- Ask Genetic Matching to optimise balance
- pick those matches that maximise balance across three covariates (e.g age, gender, social class)
- e.g. largest p value from t-tests
- Algorithm tries many alternative weights until have those that give best balance say p=0.6,0.9 and 1

Xigris for severe sepsis: subgroup with 3-5 organ failures (n=878 matched pairs) Covariate balance PSM vs IPW vs GM



Xigris for severe sepsis, subgroup with 3-5 organ failures Cost-effectiveness results

	Using subgroup specific PS mean (95% CI)*		
	Inc cost £	Inc QALY	INB**
Genetic Matching	19,948 (17,610 to 22,286)	1.28 (0.86 to 1.70)	5,690 (-2,543 to 13,924)
IPW	19,023 (15,636 to 22,102)	0.542 (-0.66 to 1.55)	-8,175 (-31,787 to 11,845)
Pscore matching	19,384 (17,696 to 21,071)	0.98 (0.65 to 1.33)	391 (-6,350 to 7,133)

*Non-parametric bootstrap CI
**INB at £20,000 per QALY

Interpretation and next talks

- GM better balance observed covariates than PS methods
- Similar results when applying other methods
- Size of control pool, choice of covariates and initial overlap is important
  - Line of therapy
  - Time
- No unobserved confounding must be fully discussed

Instrumental variable methods unobserved confounding and heterogeneity

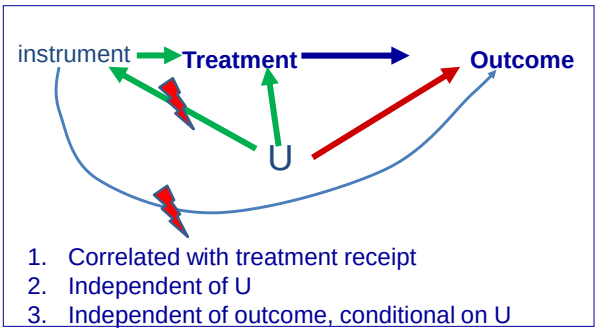
- Major problem: only some covariates are observed
- Treatment selection on unobserved effect modifiers
- “Essential heterogeneity”, Heckman & Vytlacil, 1999
- For example, clinicians use knowledge about pre-admission health status to decide whether to transfer patient to ICU
- Instrumental variable (IV) estimands limited relevance
- Person-centered Treatment (PeT) effect (Basu 2012)

IV study: evaluation of ICU admission: spotlight study

- 'Deteriorating ward' patients
- What is the effect of ICU admission on mortality?
- Precedent studies biased estimates inadequate design and analysis
- Spotlight study, 13,011 patients (48 hospitals) assessed for ICU admission
- ICU transfer vs not (general medical)
- Regression analysis, found ICU transfer led to increased mortality
- IV: variation in number of available ICU beds at time of assessment

Harris S et al. *Lancet* 2015;386 Suppl 1: S40.

All relies on valid IV



Average effects of ICU transfer versus not

Instrumental variable estimates
absolute risk differences

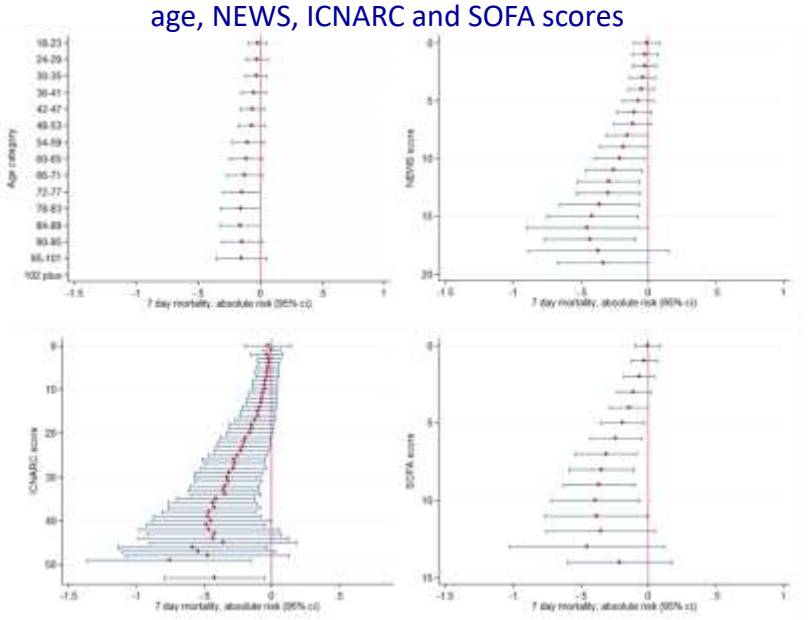
	ICU effectiveness mean (95% CI)
7 day mortality	**-11.7% (-25% to 1.5%)**
28 day mortality	**-4.9% (-26.4% to 16.7%)**
90 day mortality	**-4.7% (-28.5% to 19.2%)**

- Non-parametric bootstrap CI

**INB at £20,000 per QALY

PeT Effects by subgroup:

age, NEWS, ICNARC and SOFA scores



Interpretation and further research

- IV approach allowed for unobserved confounding, and heterogeneity
- Developed models predict magnitude of treatment effects
- Warrants extensions other settings such as large administrative data
- Good design is key
  - Minimise observed confounding
  - Collection of IV
- Time is ripe for observational methods and continued testing and refinement as required as they enter mainstream decision-making

References

Diamond, A. & Sekhon, J. S. 2013. Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. Review of Economics and Statistics. 95(3): 932-945

Sekhon, J. S. 2011. Matching: multivariate and propensity score matching with automated balance search. Journal of Statistical Software. 42(7)

Sekhon, J. S. & Grieve, R. 2011. A Matching Method for Improving Covariate Balance in Cost-Effectiveness Analyses. Health Economics 21(6):695-714

Kreif N et al. (2014). Regression-adjusted matching and double-robust methods for estimating average treatment effects in health economic evaluation. Health Services and Outcomes Research Methodology 13 (2-4), 174-202, 2013.

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.

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