Adjusting for unmeasured confounding in non-randomised longitudinal studies - a methodological review

Adam J. Streeter

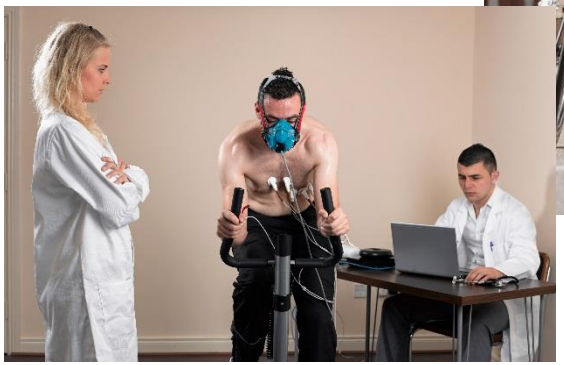
University of Exeter Medical School
Athena SWAN Silver Award
Plymouth University Peninsula Schools of Medicine & Dentistry
Athena SWAN Silver Award

	Trial data	Health records
Cost	££££	Usually cheaper
Size	Restricted by cost	Potentially huge!
Follow up	Restricted by £/y	Usually longer

Uncontrolled, but wide variety of risk groups

Generalisability

Controlled, but restricted sample



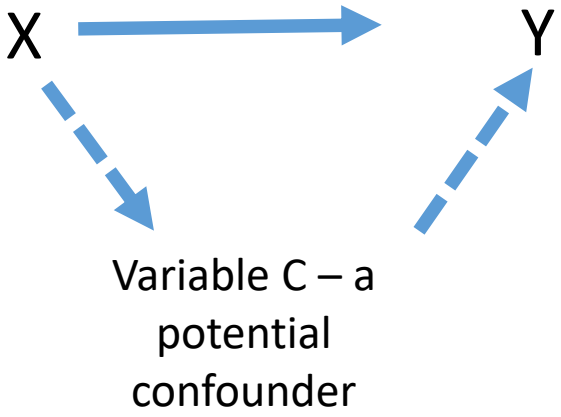
Two levels of evidence

- Causal:

  Does X cause Y?

- Association:

  We can see X is correlated with Y, but . . .



Confounder measured?

✓ Can adjust analysis for measured confounders (covariate, balancing, propensity scores)

But,

Routinely collected data?

! Unmeasured confounders

Systematic review:

Methods used to adjust for unmeasured confounding in nonrandomised, longitudinal data

Longitudinal =
repeated observations on identifiable individuals

Patient history

Large number of observations

. . . .			Outcomes over time							Patient
					X	X			X	97735
			X	X			X	X		100633
										197065
	X		X			X			X	246748
				X			X	X	X	356419
	X		X							499871
		X			X			X		641968
	X					X			X	649874
							X		X	987911
										. . . .

Other inclusion criteria:

- Adjustment for unmeasured confounding (not through proxy variables or propensity scores)
- Explicitly acknowledged unmeasured confounding, to justify method
- Independent control arm
- Data > 1000 observations

Results

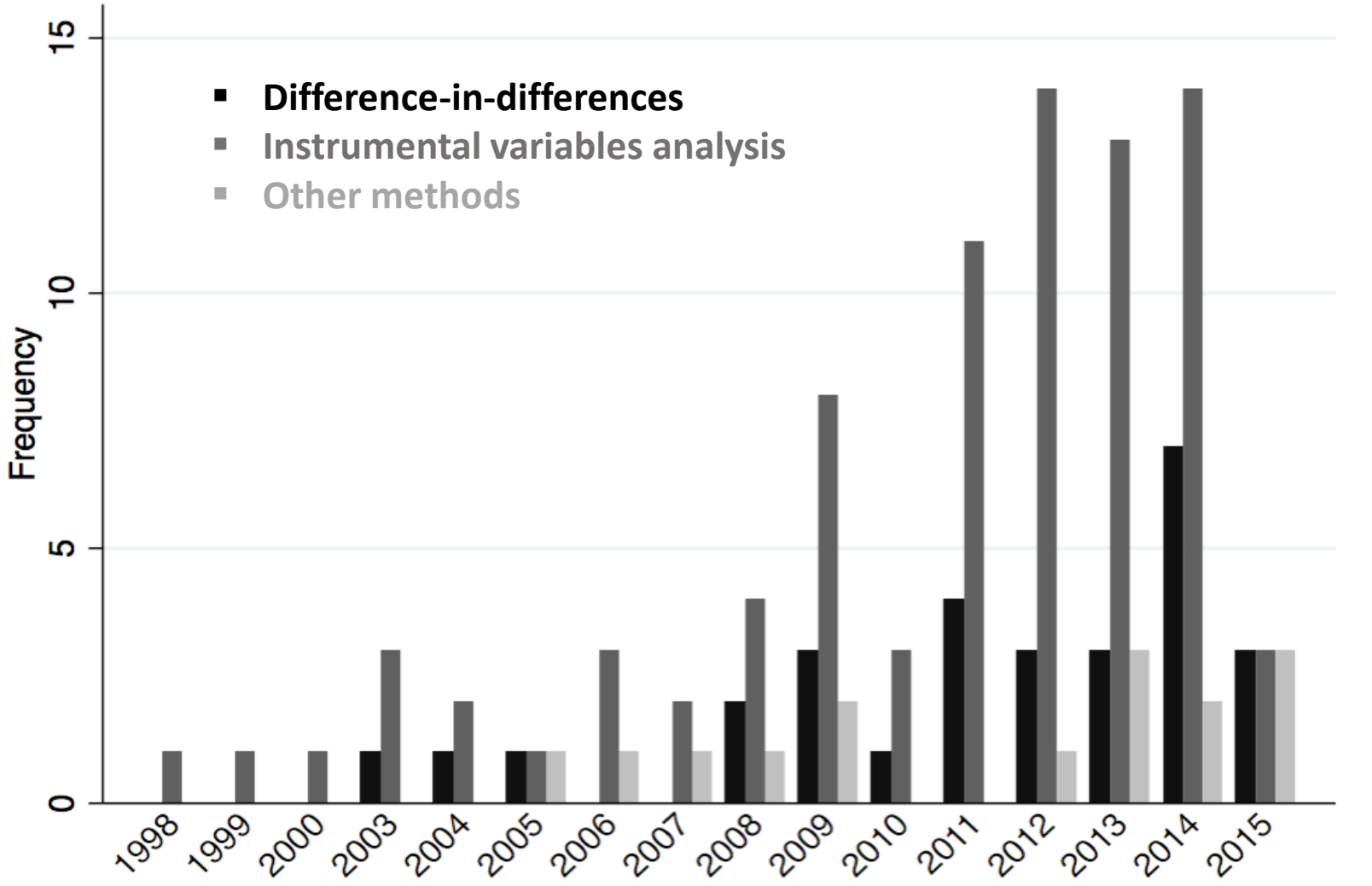
Included studies:

Titles and abstracts of 734 unique studies

Full text review of 275 studies

121 studies included in the review

Results


Difference-in-differences
Instrumental variables analysis
Other methods
Frequency
0
5
10
15
1998
1999
2000
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015

Results

84 used instrumental variable analysis (IVA)

- 36 used lagged or historical instruments.

29 used difference-in-differences (DiD) and fixed effects (FE) models

- 5 combined IVA with DiD or FE to try to mitigate for time-dependent confounding.

Other: prior event rate ratio adjustment, regression discontinuity nested within pre-post studies, propensity score calibration, perturbation analysis and negative control outcomes.

Conclusions

- Well-established methods from econometrics (DiD and IVA) commonly address unmeasured confounding
- Information from longitudinal dimension not always fully utilised
- Slow uptake & knowledge translation of statistical innovations – consistent with prior research on dissemination[1]

1. Pullenayegum EM, Platt RW, Barwick M, Feldman BM, Offringa M, Thabane L. **Knowledge translation in biostatistics: a survey of current practices, preferences, and barriers to the dissemination and uptake of new statistical methods**. *Stat Med*. 2016.

Conclusions

Roger's Diffusion of Innovations model for adoption of novel methodologies[2]:

1. Clearly describe methods using foundation principles
2. Compare results to established methods
3. Provide sample data, code or calculation examples
4. **Early communication support and testing**
5. **Provide methodological and reporting guidance**

2. Cadarette SM, Ban JK, Consiglio GP, Black CD, Dubins D, Marin A, Tadrous M. **Diffusion of Innovations model helps interpret the comparative uptake of two methodological innovations: co-authorship network analysis and recommendations for the integration of novel methods in practice**. *J Clin Epidemiol*. 2016.

Conclusions

Evaluate relative performance of emerging methods in a range of applied contexts and settings.

Thank you

Streeter AJ[a,b], Lin NX[c], Crathorne L[b], Haasova M[b], Hyde C[b], Melzer D[b], Henley W[b].

Adjusting for unmeasured confounding in non-randomised longitudinal studies: a methodological review.
Journal of Clinical Epidemiology 2017.

doi:10.1016/j.jclinepi.2017.04.022

a - Plymouth University Peninsula School of Medicine & Dentistry, Plymouth.
b - University of Exeter Medical School, Exeter
c - Mathematics, physics & electrical engineering, Northumbria University, Newcastle-upon-Tyne

Results

IV type	Explanation/ Example	No. of papers				Total frequency
Mendelian	Genetic characteristics :Single nucleotide polymorphisms	11				11
Geographic	Differential distance between patient's postcode and nearest health facility	19	1		1	20
Time	Time-based characteristic of treatment such as date of therapy	6		2		10
Historical	Usually prescribing preference of physician or facility based on historical records of previously administered therapies	31				34
Lagged	Previous therapy or outcome of patient	6				6
Randomisation	Original randomisation	1				1
Other	Characteristics of individual e.g: age of patient, weight of offspring	8				8

Method	Description	Obstacles to implementation	Frequency of methods
Instrumental variable analysis (IVA)	Upon identification of a suitably strong instrument, the influence of bias may be reduced through post-hoc randomisation. The instrumental variable should be highly determinant of the intervention or treatment received, while satisfying the exclusion assumption of being independent of the outcome other than through the treatment (Wright 1928; Angrist 1991).	In practice, finding an instrument with a sufficiently strong treatment association is a stumbling block in many analyses (Bound, Jaeger, and Baker 1995; Baser 2009). Association of the instrument with the outcome exclusively through the treatment is an untestable assumption, particularly if an indirect association exists through an unmeasured covariate.	79
Difference-in-differences (DiD)	A biased effect estimate between two treatment groups may be corrected by the same estimates from a treatment-free period prior to the exposure, which should be a measure of the confounding bias contributed to the treatment effect (Ashenfelter and Card 1984). Aggregated at the treatment group level, this is operationalised in regression as a period-treatment interaction. At an individual level, demeaning, first-differencing or dummy variables for each individual may yield bias-free fixed effects, contingent on assumptions.	The method is contingent on the availability of repeated outcomes in both periods and invokes a time-invariant confounding assumption: that the confounding bias as captured by the estimated treatment effect in a treatment-free period prior to exposure is constant through to the study period.	24
Prior event rate ratio (PERR)	Analogous to the DiD method for time-to-event or rate data, a biased estimate of the hazard ratio or the incidence rate ratio is adjusted through its ratio with that from a treatment-free prior period (Tannen et al. 2008).	As with the assumption for DiD, repeatable outcomes and a constancy of the unmeasured confounding bias is required across both periods, before and after the exposure. Prior event occurrence should not influence the likelihood of future treatment.	5
Fixed effects instrumental variable analysis (FE IVA)	IVA may be applied to DiD estimation to mitigate for second-order endogeneity: the time-varying part of the bias that may not have been adjusted for by DiD.	Assumptions of IVA apply	5
Dynamic panel model, or Instrumental variable - generalised method of moments (IV-GMM)	Lagged observations of the confounded (endogenous) explanatory variable are introduced in a first-differences fixed effects analysis so that the differences of the lags become the instrumental variables in a generalised method of moments estimation.	Assumptions of IVA apply. Here the differenced lags should not be correlated with the differences in the error terms.	2
Regression discontinuity (RD)	RD is a design for analysis based on a treatment assignment determined by a cut-off applied to a continuous variable that is preferably measured with some random noise (as many clinical tests may be). The outcome can then be modelled on treatment for individuals within a certain interval from the cut-off of the assignment variable to ensure exchangeability between individuals for robust causal inference (Thistlethwaite and Campbell 1960)	Where assignment is not sharply determined by the cut-off, an increase in the probability of treatment may be observed leading to a "fuzzy" version of RD. Continuity in the assignment variable is assumed, otherwise manipulation of assignment and reverse causality may be suspected. Assignment should be locally random around the cut-off and makes the weak assumption that no unobserved covariates are discontinuous around the assignment cut-off.	3
Propensity score calibration (PSC)	PSC adjusts for residual confounding in the error-prone main dataset by importing information about the unmeasured confounders from a smaller, external “gold-standard” dataset (Stürmer et al. 2005). Analysis in the main dataset is adjusted using a single dimension propensity score of the measured corrected for unmeasured confounding by regression calibration against the gold-standard propensity score.	Successful adjustment is wholly dependent on the availability of another dataset containing the exposure variable and error-free predictor, with individuals that are relevant enough to those in the main dataset and under similar enough conditions to assure sufficient overlap between the two datasets.	3
Perturbation testing/analysis (PT/PA)	This data mining approach aims to mitigate for unmeasured confounding by adjusting for many measured variables that are weakly associated with the unobserved confounding variables (Lee 2014). Simulation in the single reviewed example demonstrated this may require 100's, if not 1000's of perturbation variables (PV).	This requires a very highly dimensional dataset, which may ultimately obviate the need for indirect adjustment if the most or all of the confounders are captured. Simulation demonstrated the bias may be exaggerated if a confounder is inadvertently identified as a PV, requiring many more true PVs to correct the bias. The number of PVs may exceed the available degrees of freedom necessitating clustering.	1
Negative control outcome / exposure (NCO/NCE)	A negative controlis causally related to measured and unmeasured confounders affecting the exposure and main outcome, but not directly causally related to exposure and outcome themselves. As such, the negative control may be used to detect confounding bias in the main study, and potentially to indirectly adjust for this (Richardson et al. 2014)	This assumes that the effect of the unmeasured confounders on the main outcome is similar to that affecting the negative control.	1