

Epidemiological issues in the design and conduct of natural experiments

barney.reeves@bristol.ac.uk

Outline

- Perspective
- Validity and applicability
- Risk of bias
- Systematic review examples
- Ways forward

Perspective

- Co-convenor: Cochrane Non-Randomised Studies Methods Group)
- “Chapter 13” in Cochrane Handbook: Including Non-Randomised studies in Systematic Reviews of Interventions’
 - Make sure research question is a priority
 - Review evidence systematically (very much more difficult and time consuming)
 - Display / describe evidence in an unbiased manner (forest plot)
 - Pool data only in exceptional circumstances

Perspective

What is the debate about?

- No debate for some outcomes:
 - “Unintended” (harms), rare (but important) outcomes, typically serious adverse events
 - Outcomes arising a long time after the intervention
- Major debate about “intended” outcomes / benefits
- Major debate about
 - The feasibility of randomised controlled trials (RCTs), and
 - Whether RCTs can be done sufficiently real-world settings to inform policy decisions

Validity and applicability

		Risk of bias	
		Low	High
Applicability	Good	✓	?
	Poor	?	x

Validity and applicability

How to have good applicability and low risk of bias?

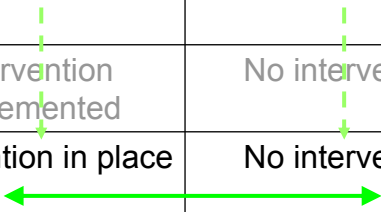
- Pragmatic RCT, or
- Natural experiment with low risk of bias

Which natural experiments have a low risk of bias?

- Population-based interventions often target subjects in 'clusters' – complex designs
 - "[Cohort] controlled before-and-after studies" (CBA)
 - "[Controlled] interrupted time series" ([C]ITS)
- Other prospective studies (?)

Validity and applicability: CBA

Study phase	“Intervention” (sites = n_I)	“Control” (sites = n_C)
Time _{before}	No intervention	No intervention
Time _{impl}	Intervention implemented	No intervention
Time _{after}	Intervention in place	No intervention



“Effect” = ‘change’ in Intervention vs. ‘change’ in Control

Validity and applicability: CITS

Study phase	“Intervention” (sites $\geq 1_I$)	“Control” (sites $\geq 1_C$)
Time 1	No intervention	No intervention
Time 2	No intervention	No intervention
...	No intervention	No intervention
Time ≥ 3	No intervention	No intervention
Time _{impl}	Intervention implemented	No intervention
Time n+1	Intervention in place	No intervention
Time n+2	Intervention in place	No intervention
...	Intervention in place	No intervention
Time n+ ≥ 3	Intervention in place	No intervention

Validity and applicability: ITS

Study phase	"Intervention" (sites $\geq 1_i$)	
Time 1	No intervention	
Time 2	No intervention	
...	No intervention	
Time ≥ 3	No intervention	
Time _{impl}	Intervention implemented	
Time n+1	Intervention in place	
Time n+2	Intervention in place	
...	Intervention in place	
Time n+ ≥ 3	Intervention in place	

Validity and applicability

Effect estimates from RCTs & NRS:

- 8 methodological reviews comparing effect estimates from RCTs and non-randomised/observational studies (majority NOT population-based interventions)
- No assessment of 'pragmatism'

Validity and applicability

Appraisal of 8 reviews:

- Identification of studies:
“The seven reviews in medical areas were each only based on a subset of known comparisons of randomised and non-randomised evidence.” (p.18)
- Similarity of interventions & outcomes in RCTs and NRS:
Issue noted by 6/8; rigorous attempts to address made by only 2/8.

Deeks et al, HTA 2003

Validity and applicability

Appraisal of 8 reviews:

- Sensible criteria to determine difference/equivalence:
“The manner in which results were judged to be ‘equivalent’ or ‘discrepant’ varied widely between reviews and influenced the conclusions that were drawn.” (p.19)
- Similar methodology apart from method of allocation
“Importantly, the possible biases of non-randomised studies vary with study design. [Except for one ...] RCTs were compared with non-randomised studies of a mixture of designs.” (p.19)

Deeks et al, HTA 2003

Validity and applicability

Conclusion of meta-review:

- *“The conclusions of the eight reviews are divergent, and all have weaknesses ... in some circumstances, the results of RCTs and non-randomised studies differ, but it cannot be proved that differences are not due to other confounding factors.”*
- Similarly, cannot prove that ‘equivalences’ are not due to other confounding factors.
- One factor leading to BRANDO collaboration.

Deeks et al, HTA 2003

Risk of bias

Bias	Description	RCTs	NRS
Selection bias	Imbalance between groups, “case-mix” problem	No	Yes
Performance bias	Co-intv’ns given differentially because intervention known	If not blinded	If not blinded
Detection bias	Outcome assessment biased because intervention known	If not blinded	If not blinded
Attrition bias	[Differential] Loss to follow-up, excluded from analysis	Possibly	Probably
Selective reporting bias	Report analyses or outcomes selectively (e.g. $p < 0.05$)	Possibly	Probably

Risk of bias (RoB) assessment tool

Assessment item	Judgement (“high”, “low” or “unclear” risk of bias)	Description
Sequence generation	RoB “high” for <u>all</u> non-randomised studies (NRS)	[Reviewer completes]
Allocation concealment	RoB “high” for <u>most</u> non-randomised studies (NRS)	[Reviewer completes]
Blinding (<u>per outcome</u>)	RoB “high” for <u>most</u> non-randomised studies (NRS)	[Reviewer completes]
Incomplete outcome data (<u>per outcome</u>)	RoB “high” for <u>some</u> non-randomised studies (NRS)	[Reviewer completes]
Free of selective reporting	RoB “high” for <u>most</u> non-randomised studies (NRS)	[Reviewer completes]

Risk of bias

Selection bias / confounding:

- Perceived to be the ‘key’ issue
- Difficulty in characterising confounding factors
- Statistical methods unable to control completely

Risk of bias

Selective reporting:

- Current focus on selective reporting of outcomes
- Investigated by comparing protocols with published reports of RCTs [Dwan et al., 2008 PLoS one]
 - “... statistically significant outcomes had a higher odds of being fully reported compared to nonsignificant outcomes (range of odds ratios: 2.2 to 4.7).”
- Substantial problem in RCTs (even cancer)

Trial ID (author, date of publication)	Review primary outcome Overall survival	Review Outcomes					Other Trial Outcomes	
		Event- free survival	Overall remission rate	Relapse rate	Toxicity & adverse events	Quality of life	Relaps e site	Time to relapse
Anderson 1983	x	x	x	x	x	x	x	x
Brecher 1997	✓	✓	✓	x	✓	x	x	x
Cairo 2003a	x	○	x	x	x	x	x	x
Magrath 1973	x	x	x	✓	x	x	x	✓
Magrath 1976	✓	x	x	✓	x	x	✓	✓
Neequaye 1990	✓	x	x	✓	x	x	x	✓
Olweny 1976	✓	x	✓	✓	x	x	x	x
Olweny 1977	✓	x	x	✓	✓	x	x	x
Patte 1991	✓	✓	○	x	✓	x	x	x
Sullivan 1991	x	✓	x	x	x	x	x	x
Ziegler 1971	x	x	x	✓	x	x	x	x
Ziegler 1972a	✓	x	x	✓	x	x	○	○

Risk of bias

Selective reporting:

- Current focus on selective reporting of outcomes
- Investigated by comparing protocols with published reports of RCTs [Dwan et al., 2008 PLoS one]
 - “... statistically significant outcomes had a higher odds of being fully reported compared to nonsignificant outcomes (range of odds ratios: 2.2 to 4.7).”
- How big a problem is this likely to be in NRS?
 - ‘Selection’ is potentially a problem for conduct, analysis and reporting
 - Selection is minimised by having a detailed protocol

Risk of bias

Methodological research on bias:

- To estimate the direction and magnitude of bias and the uncertainty bias introduces
 - Empirically (BRANDO)
 - Using expert judgements (BAMMOS)
- Additional uncertainty tends to lead to inconclusive findings
- Magnitude, direction and variability of bias in different circumstances not often directly observed?

Systematic review examples

Speed enforcement detection devices for preventing road traffic injuries. Wilson et al. 2006, CDSR Issue 2: CD004607

Number of CBAs / ITS	22/4
Number of CBA sites: >7&>7/ other/ 1&1/ not rep'd	5 / 8 / 3 / 6
Interventions	Varied: fixed/mobile cameras, overt/covert, combinations
Outcomes	Varied: speed, crashes, combinations
Settings	Varied: urban, semi-rural, rural, mixed, different road types/speed limits
Risk of bias (low/fair/high)	6 / 6 / 10
Analyses: - primary - review	Varied ++ Simple, e.g. % change, change ratio
Synthesis	Narrative, no forest plot

Population-based interventions to prevent of fall-related injuries in older people. McClure et al. 2005, CDSR Issue 1: CD004441

Number of CBAs / ITS	6/0
Number of CBA sites: >7/ other/ 1&1/ not rep'd	0 / 2 / 4 / 0
Interventions	Varied: home visits, hazard protection, physical activity, etc., combinations
Outcomes	Varied: fall-related injuries, fractures
Settings	Varied: country, urban, semi-rural, rural, mixed
Risk of bias (low/fair/high)	1 / 3 / 2
Analyses: - primary - review	Varied ++ Simple, e.g. % change, change ratio
Synthesis	Narrative, no forest plot

Systematic review examples

Conclusions:

- *“Despite the methodological limitations of the studies reviewed, the consistency of reported reductions in ...*
- *“... speed and crash outcomes across all studies suggest that SEDs are a promising intervention for reducing the number of road traffic injuries and deaths. More studies of a scientifically rigorous nature are necessary to provide a stronger evidence base ...”*
- *“... fall-related injuries across all programmes support the preliminary claim that the population-based approach to the prevention of fall-related injury is effective and can form the basis of public health practice. Randomised, multiple community trials of population-based interventions are indicated to increase the level of evidence in support of the population-based approach.”*

Ways forward

- Improve natural experiments [and pragmatic RCTs]
- Discourage the view that natural experiments are a “quick and easy/dirty” option
- Involve relevant methodologists at the outset (cf. Clinical Trial Units and RCTs)
- Prioritise research questions for which evaluations are required – and invest funding accordingly
- Weigh up benefits and harms of uncertain decision now vs. waiting for better evidence

Ways forward

Suggestions improving natural experiments:

- Establish guidelines for CBAs / ITS
- Pre-specified protocol (cf. RCT) defining:
 - Population, intervention, comparator and outcomes
 - Potential confounders and how they will be measured
 - Susceptibility of outcome to detection bias
 - Strategy for dealing with incomplete outcome data
 - Analysis plan
- Register studies with (a) prospective data collection or (b) which require access to data through ‘guardians’ (date-stamp availability of data)