



Federal Institute
for Drugs
and Medical Devices



Use of synthetic controls A rejoinder from regulatory statistics

Norbert Benda

Disclaimer:

Views expressed in this presentation are the author's personal views and not necessarily the views of BfArM

Randomized trials in areas where real-world data are widely available

American Economic Journal: Applied Economics 2015, 7(1): 1–21
<http://dx.doi.org/10.1257/app.20140287>

Six Randomized Evaluations of Microcredit: Introduction and Further Steps[†]

By ABHIJIT BANERJEE, DEAN KARLAN, AND JONATHAN ZINMAN*

Causal evidence on microcredit impacts informs theory, practice, and debates about its effectiveness as a development tool. The six randomized evaluations in this volume use a variety of sampling, data collection, experimental design, and econometric strategies to identify causal effects of expanded access to microcredit on borrowers and/or communities. These methods are deployed across an impressive range of locations—six countries on four continents, urban and rural areas—borrower characteristics, loan characteristics, and lender characteristics. Summarizing and interpreting results across studies, we note a consistent pattern of modestly positive, but not transformative, effects. We also discuss directions for future research. (JEL D14, G21, I38, O15, O16, P34, P36)



Why are political or economic decisions always highly controversial?

- of course: different interests
- but also because evidence is highly controversial
 - despite lots of real world data
- however, real world data
 - do not respond to the question about best decision
 - do not address the counterfactual

Drug approval:

- drug approval should also addresses the question on
 - the treatment decision
 - is it better for a specific patient to take
 - A or B , A or nothing ?
 - directly addressed by an experiment, i.e. RCT, not by observation

Regulators question

What is the question that I am interested in?

- How can I predict the impact of
 - the treatment that is prescribed ? = the decision that is taken ?
- Counterfactual question
 - doing X vs not doing X

Synthetic controls – external data – real-world data

Definitions not always clear –

- external data
 - = external to randomized clinical trial in the relevant population (relevant for the intended indication)
- real-world data
 - = data obtained in an observational “real” setting
 - non-experimental data
- real-world data \subset external data

External data for synthetic controls

- data generated outside randomized controlled trials
 - observational studies
 - registries
 - etc.
- data generated in different settings/populations
 - different populations
 - different underlying soc
 - etc.

Reasons given to use of external (or RWD) data

1. RCT with reasonable size considered to be difficult, unfeasible, or unethical
 - e.g.
 - orphan diseases
 - paediatric applications
 - balancing medical need vs robustness of conclusions
 - avoid/minimize placebo control
2. External validity supposed to be improved by “real-world” data
 - RCT setting considered to lack external validity
 - support RCTs conclusions by additional information

Avoid placebo treatment in clinical trials ?

- ethical justification of a clinical trial usually based on **equipoise**
 - unknown whether B/R is positive as compared to placebo
- do you question equipoise ?
 - or are you just overoptimistic about your drug ?
- later properly controlled trial may never be possible anymore
 - prevention of evidence may be highly unethical

Rare diseases:

- rare diseases may prevent from large studies
 - but not from (small) randomized trials

What should we ask when external data are used?

- Are external data needed?
 - additional evidence may be weak
- How have they been selected?
 - kind of pre-specification
 - selection bias: selection of patients – selection of studies
- What are the underlying assumption in a combined evaluation?
 - robustness of model assumptions used
- What is the potential bias?
 - additional non-randomized evidence:
 - causality to treatment, bias in absolute response
 - additional randomized evidence in different settings/populations:
 - treatment-by-population interaction

Required evidence in drug approval

- internal validity
 - comparison and randomization ensures causality in the presence of counterfactuals
 - pre-specification and type-1 error control limits false positive decisions
 - RCT ensures validity within study setting
- external validity
 - generalizability on “real world”
 - potential issues often mentioned:
 - selected centres, investigators, additional care, patient’s expectations, narrow inclusion criteria, etc.
- focus on internal validity in RCTs

Internal and external validity

- RCT: internal validity
 - randomisation ensures causality
 - pre-specification controls false positive decisions
- external validity in RCTs
 - usually: “WYSIWYG principle”: “what you study is what you get”
 - indication informed by study design (inclusion criteria, etc.)
 - assumption made:
 - absolute effect may change, but relative effect constant
 - only extrapolation of treatment difference to real world required
 - treatment-by-design interaction usually less relevant than lack of internal validity in non-interventional studies or comparison to historical controls

Issues with real-world data

- Real-world data, non-interventional studies
 - no randomization, no direct comparison
 - indirect comparison only
 - no proof of causality
 - difficult or incomplete adjustment for covariates
 - selection of “unplanned” data
 - potential selection bias
 - combining different sources of evidence
 - combining randomized studies, observational studies, registries etc.
 - relying on (unverifiable) model assumptions
 - robustness unclear

Type-1 error control

limit false positive decisions

- control of type-1 error conditional on given (fixed) external data only does not reflect properly the probability of a false positive decision
- type-1 error control should (ideally) include generation of external data
 - may be difficult due to lack of pre-specified protocol
 - may be difficult due to complex model selection
- type-1 error $\rightarrow 1$ for increasing sample size if bias > 0

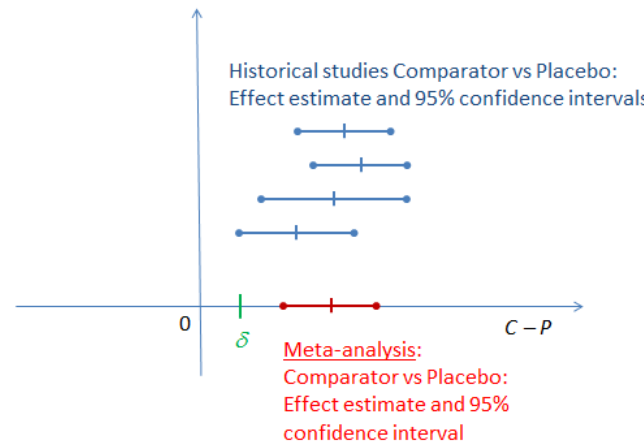
Bias

- can we really quantify reliably bias / difference to external setting ?
- can we really adjust reliably ?
- do we trust the underlying assumption ?
- lot's of efforts made, but how do we verify ?
 - post-hoc quantification/justification in a specific situation does not necessarily convince me for a future trial
- account for remaining uncertainty could mean :
 - test vs threshold instead of 0 – still to quantify a reasonable threshold

Analogy to derivation of a non-inferiority margin ?

- historical data used to define non-inferiority margin
 - putative placebo comparison
 - here: usual approach is not to analyse both source jointly
 - but to account for remaining uncertainty

Non-inferiority margin: Statistical justification



e.g. $\delta = \frac{1}{2}$ of lower limit of the 95% confidence interval of $C - P$

Summary

- “real-world” data describe what happens but not which treatment decision is best – and so do synthetic controls
 - data only vs experiments
- use of synthetic controls
 - may be used as additional evidence to (self-standing) RCTs
 - if you know that placebo response would be 0
 - when do you know ?
 - is this the relevant endpoint ? – what about survival ?
 - if discussed in case of limited options to generate randomized data
 - small RCTs may still be possible
 - validity may still be difficult to justify
 - results to be qualified, may require stricter requirements w.r.t. the null hypothesis
 - usually based on assumptions difficult to verify
 - difficult to control type-1 error / false positive decisions