HEIDELBERG
UNIVERSITY
HOSPITAL

Synthetic controls – what do we need and how far can we go?

Rejoinder

Meinhard Kieser

Institute of Medical Biometry and Informatics (IMBI)

University of Heidelberg

imbi

UK HD

Some recent headlines – facts, prophecies, fictions?

"…[the Flatiron acquisition is] *a showcase in the future: trial designs where the currently used ones look like stagecoaches versus ICE trains*".

(Gross, 14.05.2018)

Some recent headlines – first remark

It was really Kaspar who I recently saw riding through Basel ...

Some recent headlines – second remark

Sometimes a stagecoach is more comfortable than an ICE train …

Outline

- synthetic controls – why and when?
- the crucial role of standards
- topics of current and future methodological research
- conclusions

Synthetic controls – why and when?

Synthetic controls – why and when?

synthetic controls are used

- in situations when RCTs are not feasible ✓
- "maybe broader" (Rufibach & Burger): due to disadvantages of RCTs (and (assumed) advantages of RWD trials):

  ⇒What are the advantages of RWD trials and can we quantify them?

  ⇒weighting may depend on area of application

Synthetic controls – areas of applications

synthetic controls are used for

- internal decision making
- proof-of-concept
- regulatory approval

⇒ different aims may be reflected by different requirements, challenges, and methods

⇒ different aims may lead to different weighting of (dis)advantages

⇒ different aims may require different standards (see below)

Mentioned disadvantages of RCTs – potential motivations for “maybe broader”

- high costs: What are the actual costs of RWD trials?

  “*Under the terms of the agreement, the transaction value for the acquisition of Flatiron Health was USD 1.9 billion…*”

  (Roche Media Release, 06.04.2018)

Mentioned disadvantages of RCTs – potential motivations for “maybe broader”

- high costs
- long time to complete: What is the actual timeframe of RWD trials?

  fewer patients to be recruited

  but:

  additional tasks to be performed

Mentioned disadvantages of RCTs – potential motivations for “maybe broader”

- high costs
- long time to complete
- comparison are possible only against few alternative treatments:
  Yes, but RWD trials are helpful as alternative only if they lead to “robust conclusions” (Salmonson acc. to Brookland) / “do not alter relevantly the approval bar” (Rufibach & Burger) (see below)

Mentioned disadvantages of RCTs – potential motivations for “maybe broader”

- high costs
- long time to complete
- comparison are possible only against few alternative treatments:
- eligibility criteria:
  - impractical
  - leading to non-representative populations

Can the RCT culture be changed?

Mentioned disadvantages of RCTs – potential motivations for “maybe broader”

- high costs
- long time to complete
- comparison are possible only against few alternative treatments:
- eligibility criteria:
  - impractical
  - leading to non-representative populations
- do not make use of available data / unnecessary placebo exposure:
  Yes, but RWD trials are helpful as alternative only if they lead to “robust conclusions” (Salmonson acc. to Brookland) / “do not alter relevantly the approval bar” (Rufibach & Burger) (see below)

Mentioned disadvantages of RCTs – potential motivations for “maybe broader”

- high costs
- long time to complete
- comparison are possible only against few alternative treatments:
- eligibility criteria:
  - impractical
  - leading to non-representative populations
- do not make use of available data / unnecessary placebo exposure
- RCT effects are effects under ideal conditions and exaggerate the RW effect:

Is this really true? “… *there is little evidence for significant effect estimate differences between observational studies and RCTs*…” (Anglemyer et al., 2014)

Mentioned disadvantages of RCTs – potential motivations for “maybe broader”

- high costs
- long time to complete
- comparison are possible only against few alternative treatments:
- eligibility criteria
  - impractical
  - leading to non-representative populations
- do not make use of available data / unnecessary placebo exposure
- RCT effects are effects under ideal conditions and exaggerate the RW effect:

… or is rather the opposite true? “*Studies of routinely collected health data … may substantially overestimate treatment effects*” (Hemkens et al., 2016)

Mentioned disadvantages of RCTs – potential motivations for “maybe broader”

- high costs
- long time to complete
- comparison are possible only against few alternative treatments:
- eligibility criteria
  - impractical
  - leading to non-representative populations
- do not make use of available data / reduce placebo exposure
- RCT effects exaggerate the RW effect (“effect under ideal conditions”)

⇒ clarification (and, if possible, quantification) is required!

⇒ We should abandon the established gold standard only if we have substantial arguments!

imbi

The crucial role of standards

Standards

- don’t undervalue **the merits of standards**!
- for **common clinical trials**, established standards exist for…
  - data source
  - design
  - conduct
  - analysis
  - reporting
- which **standards** are available **for clinical trials with synthetic control**, which have to be created?

Standards – example: data source

- **exciting insights in processes** (Somnath Sarkar)
- **open questions:**
  - what means “high data quality”?
  - what is the benchmark?
  - how has data quality to be assessed?
  - data sources are regional:
    - global data sources required for global conclusions? (Rufibach & Burger)
    - similarity to concept of “multi-regional trials”?
  - …

Standards – reporting

"Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0 Shirley

V.Wang, PhD, ScM[1,2,*,] Sebastian Schneeweiss, MD, ScD[1,2], Marc L. Berger, MD[3], Jeffrey Brown, PhD[4], Frank de Vries, PharmD, PhD[5], Ian Douglas, PhD[6], Joshua J. Gagne, PharmD, ScD[1,2], Rosa Gini, PhD, MSc[7], Olaf Klungel, PhD[8], C. Daniel Mullins, PhD[9], Michael D. Nguyen, MD[10], Jeremy A. Rassen, ScD[11], Liam Smeeth, MSc, PhD[6], Miriam Sturkenboom, PhD, MSc[12], on behalf of the joint ISPE-ISPOR Special Task Forceon Real World Evidence in Health Care Decision Making"

Wang SV., Schneeweiss S., Berger ML., et al. (2017). Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V 1.0. *Value Health* **20**: 1009-1022.

Standards – reporting

- **key elements:**
  - agreeing on common terminology
  - specification of a (minimal) set of essential information to be reported

Table 2 – Reporting specific parameters to increase reproducibility of database studies

	Description	Example	Synonyms
A. Reporting on data source should include:			
A.1 Data provider	Data source name and name of organization that provided data.	Medicaid Analytic Extracts data covering 50 states from the Centers for Medicare and Medicaid Services.	
A.2 Data extraction date (DED)	The date (or version number) when data were extracted from the dynamic raw transactional data stream (e.g. date that the data were cut for research use by the vendor).	The source data for this research study was cut by [data vendor] on January 1st, 2017. The study included administrative claims from Jan 1st 2005 to Dec 31st 2015.	Data version, data pull

⇒ improvement of transparency, reproducibility, and credibility

⇒ required also for other aspects of RWD trials

Table2: Wang SV., Schneeweiss S., Berger ML., et al. (2017). Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V 1.0. *Value Health* **20**: 1009-1022.

Topics of current and future methodological research

Synthetic controls stimulates exciting research

- **reducing placebo exposure** (Cornelia Dunger-Baldauf et al.)
  - what astonishing results!
  - more experience desirable (also in other indications)
  - how to obtain information on safety?
- **dealing with potential biases** (Chris Harbron)
  - known methods revitalized for new context
  - applications / experiences available?

<< The future emerges before it occurs. >>

(Rainer Maria Rilke)

Topics of current and future research (cont'd)

- **matching**
  - refined methods may be required in the context of generating synthetic controls
  - example: "adaptive propensity score matching"

Topics of current and future research – adaptive propensity score matching

- **starting point:** limited data set available for control intervention (in motivating example: n=77)
- **RCT not applicable**
- **aim:** match as many control patients as possible
- **question:** how many patients are required for new intervention?
- **idea:** (re)calculate sample size at interim analysis based on observed matching rate

Topics of current and future research – adaptive propensity score matching (cont'd)

- "**naïve approach**": recalculation based on estimated matching rate

  ⇒ overoptimistic estimate of matching rate (due to high number of control patients at interim analysis)

  ⇒ too small sample size / too low power

- "**resampling CI approach**":
  - repeated random selection of control patients
  - recalculation based on lower boundary of 99% CI of matching rate

  ⇒ **very good performance** (Weber, Uhlmann & Kieser, 2019)

Topics of current and future research (con't'd)

- **new designs**

Make drug development more efficient

- Real-world data is **one option**. Other options are
  - Innovative trial designs, platform or adaptive.

(Rufibach & Burger)

⇒combine both!

⇒"matched threshold crossing design"

Topics of current and future research – matched threshold crossing design

- **starting point:** substantial dataset available for control intervention
- **heterogeneous population** (prognostic and predictive factors)
- **aim:** flexible design for decision making (or more?)
- **idea:** adaptive two-stage design with option for
  - $1:M$ matching, $M \geq 1$
  - futility stopping
  - sample size recalculation
- **very good performance**, incl. robustness against mis-specifications and control of type I error rate (if desired) (Krisam et al., 2019)

Conclusions

This word is hard to speak out in these times ….

… but the statement is essential in this context:

FOR OBJECTIVE CAUSAL INFERENCE, DESIGN TRUMPS ANALYSIS

By Donald B. Rubin

The Annals of Applied Statistics 2008, Vol. 2, No. 3, 808–840
DOI: 10.1214/08-AOAS187

- **important aspect:** pre-specification (Rufibach & Burger)

Statisticians play a key role

- **implementation** of principles and methods
- **communication** of challenges, limitations and opportunities
- **development** of methods and recommendations
- …

Statisticians play a key role (cont'd)

- **take it as an advantage and opportunity to shape the future!**
- whether you like the topic or not: **it should be a matter that is close to your heart!**

If we don't take care of it, others will do it instead of us …

Thank you for your kind attention!

References

- Anglemyer A., Horvath HT., Bero L. (2014). Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database of Systematic Reviews* **4**. Art. No.: MR000034.
- Goldsack J. (05.02.2019). *Synthetic control arms can save time and money in clinical trials*. https://www.statnews.com/2019/02/05/synthetic-control-arms-clinical-trials/. Accessed: May 07, 2019.
- Gross S. (14.05.2018). *Manhattan Love Story: Warum Roche Flatiron kauft*. https://www.handelszeitung.ch/unternehmen/manhattan-love-story-warum-roche-flatiron-kauft#. Accessed: May 07, 2019.
- Hemkens LG., Contopoulos-Ioannidis DG., Ioannidis JP. (2016): Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. *British Medical Journal* **352**: i493.

References

- Klonoff DC. (2019). The Expanding Role of Real-World Evidence Trials in Health Care Decision Making. *Journal of Diabetes Science and Technology*. DOI: 10.1177/1932296819832653
- Krisam J., Weber D., Schlenk RF., Kieser M. (2019). Enhancing single-arm phase II trials by inclusion of matched control patients – the Matched-Treshold-Crossing (MTC) design. under review at *Statistical Methods in Medical Research*.
- Pulkstenis E. (08.03.2019). *Synthetic Control Arm: The End of Placebos?* https://stories.abbvie.com/stories/synthetic-control-arm-end-placebos.htm. Accessed: May 07, 2019
- Roche Media Release (06.04.2018). *Roche completes acquisition of Flatiron Health*. https://www.roche.com/dam/jcr:4df40ca0-d500-4154-9418-83ddd700da9d/en/20180406-MR-Flatiron-Closure_en.pdf. Accessed: May 07, 2019

References

- Rubin DB. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics* **2**: 808-840.
- Shaywitz D. (11.05.2018). *Will Real World Performance Replace RCTs As Healthcare's Most Important Standard?* https://www.forbes.com/sites/davidshaywitz/2018/05/11/will-real-world-performance-replace-rcts-as-healthcares-most-important-standard/#16d01ff03557. Accessed: May 7, 2019
- Sutter S. (02.01.2019). *Pink Sheet - External Control Arms: Better Than Single-Arm Studies But No Replacement For Randomization*. https://pink.pharmaintelligence.informa.com/PS124362/External-Control-Arms-Better-Than-SingleArm-Studies-But-No-Replacement-For-Randomization. Accessed: May 07, 2019.
- Wang SV., Schneeweiss S., Berger ML., et al. (2017). Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V 1.0. *Value Health* **20**: 1009-1022.
- Weber D., Uhlmann L., Schönenberger S., Kieser M. (2019). Adaptive propensity score procedure improves matching in prospective observational trials. under review by *BMC Medical Research Methodology*.