UMC Utrecht

Synthetic controls: What do we need and how far can we go?

Discussion

Kit Roes
UMC Utrecht & CBG-MEB & EMA BSWP

The views expressed are personal views and not necessarily the views of CBG-MEB or EMA

Key reflections

- When: Can we articulate when single arm / external control evidence could be sufficient for regulatory decision making?
- How is best evidence generated (experimental design) and assessed in such a situation?
- The crucial importance of true patient registries.

When: Necessary – but not sufficient

*"In fact, **most** orphan drugs and paediatric indications submitted for regulatory approval are based on randomised controlled trials that follow generally accepted rules and guidance. Deviation from such standards is, therefore, uncommon and should only be considered when completely unavoidable and would need to be justified."*

- Rare condition
- High unmet need
- No satisfactory treatment
- *Randomisation not possible/ethical*
- No obvious control arm

- *Single arm effect "Dramatic", "Unprecedented",....*

When: Necessary – but not sufficient

Key notion to discern two objectives:

- Causality ("direct drug effect") can be inferred from the trial.
- The clinical benefit can be estimated (compared to best standard of care).

- The aim for "dramatic effect" in SAT aims to service both.

- At design stage, (clinical) effect is essentially unknown:
  - Design of SAT should still be informative if less extreme.

- These principles (causality and estimation) essentially lead to all "points of attention"
  - (from natural disease course known, to type of endpoints).

When: Necessary – but not sufficient

- Two objectives:
  - Causality (“direct drug effect”) can be inferred from the trial.
  - The clinical benefit can be estimated (compared to best standard of care).
- Designs of SATs should be able to convincingly adress both, but possibly through different endpoints, different….
- And may particularly leverage forms of external control data for the second.

When: Ethical and feasibility considerations

Just some food for thought:

Thomas Chalmers, 1975: "Randomize the first patient": "When a new dug arrives, administering this agent to patients without offering them the possible benefit, through randomisation, of being spared its potential of toxicities and/or lack of efficacy is unethical."

- All patients (experimental and control) receive best standard of care in many cases.
- Why then would it
  - not be ethical to randomize patients fortunate enough to enroll in the trial;
  - but at the same time ethical to withhold the experimental treatment to all patients not so fortunate, and mostly not by their own choice?

Ethical considerations concerning treatment allocation in drug development trials

S Senn

Stat Methods Med Res 2002; 11; 403

DOI: 10.1191/0962280202sm299ra

When: Ethical and feasibility considerations

These are probably most complex / least tangible.

- Early in clinical development randomisation more likely feasible than after initial promising results are known to the clinical community.

- Contrary to threshold crossing approach: randomise early, both decision making and the later SATs may profit.

- I agree with Thomas Salmonson, I value the possibility to randomise in these circumstances above the strict (and arbitrary) threshold we cherish.

(dramatic effect sizes do not need large sample sizes in any design)

How: Relying on external evidence

Harbron parallels surrogacy to capture additional uncertainty.

Non-inferiority trials aimed to demonstrate efficacy could similarly be a parallel to learn from to build framework.

- For the margin similar considerations and assumptions
- Conservatism included in assessment

(But still randomised, and even then maybe not our favorite)

How: Experimental design and modeling

Type 1 Error

- Imaginary quantity.
- Associated with “decision procedure”, based on the design and a specific statistical model.
- Which we (have to) agree to be plausible before the data are collected.

Control

- Has brought us many good things for confirmatory trials.
  - A rational approach to sample size choice
  - Careful pre-planning of the whole trial (good experimental design)
  - No “free lunches”
  - Clear threshold for proceeding to secondary assessment
  - At least some control of regulatory error rate
  - Level playing field
- **In settings with sufficient prior data and knowledge.**

How: Experimental design and modeling

Q -> Design -> *Data* **<->** *Analysis* -> Conclusion

Design:

- The external data may be richer than "just" to distill a control group. Could we move to a more DoE approach (incl modeling) to leverage the richness?

Data <-> Analysis in this setting:

- Agreeing a priori on a plausible model - and sticking to it because of T1E - may lead to larger errors than making sure that statistical inference is based on a model that is adequately supported by the data.

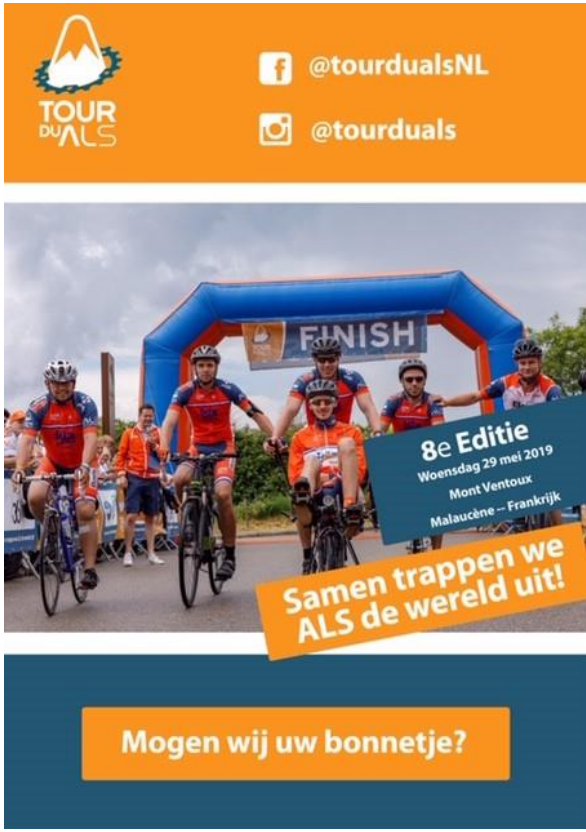
- We need a broader approach to quantifying "error" (characteristics of the decision procedure), to include the model building step.

- This is not unique to "exceptional circumstances": estimands, new high volume data, new treatment modalities,… will require the same.

RWD: Randomisation is not the problem

1. Systematic review of RCTs in ALS (2000 - 2017)
   - Placebo-controlled
   - Clinical endpoint
   - Single agent

2. Incidence-cohort UMC Utrecht (N = 2904)
   - 2006 - 2016
   - Survival & functional (ALSFRS-R) data

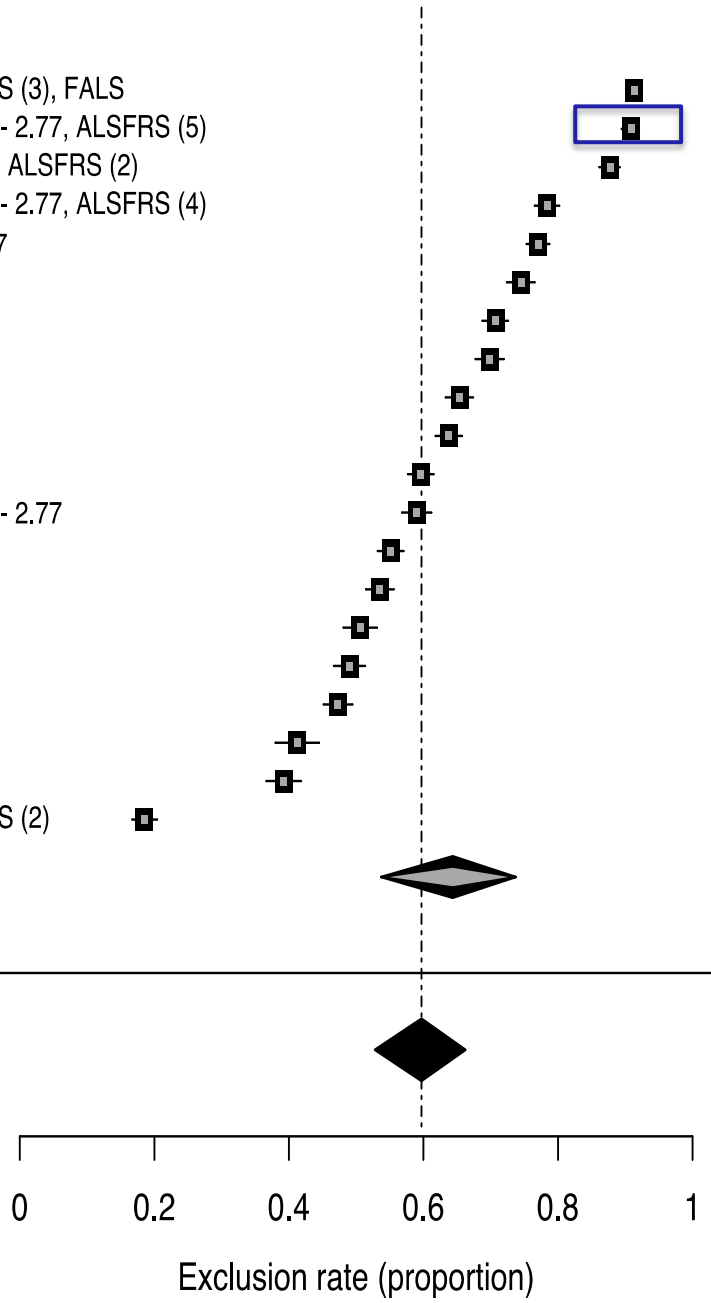


Period: 2010 - 2017

Acetyl-L-carnitine	2013	II	82	Def, Prob (LS)	≥ 80%	6 - 24	40 - 70	ALSFRS (3), FALS
Edaravone	2017	III	137	Def, Prob	≥ 80%	0 - 24	20 - 75	LI 0.36 - 2.77, ALSFRS (5)
Memantine	2010	II	63	Def, Prob (LS)	≥ 60%	0 - 36	18 - 75	LI >1.0, ALSFRS (2)
Edaravone	2014	II	205	Def, Prob (LS)	≥ 70%	0 - 36	20 - 75	LI 0.36 - 2.77, ALSFRS (4)
TUDCA	2016	II	29	Def, Prob	≥ 75%	0 - 18	18 - 75	LI <2.77
Pioglitazone	2012	II	218	All	50%-95%	6 - 36	≥ 18	
Erythropoietin	2015	III	200	Def, Prob (LS)	≥ 70%	0 - 18	18 - 75	FALS
Olesoxime	2014	III	512	Def, Prob (LS)	≥ 70%	6 - 36	18 - 80	
Lithium	2012	II	133	Def, Prob (LS)	≥ 70%	6 - 36	18 - 85	-
Flecainide	2015	II	54	Def, Prob	≥ 50%	0 - 60	18 - 75	
NP001	2015	II	136	Def, Prob	≥ 70%	0 - 36	21 - 80	-
Bromocriptine	2016	II	36	All	≥ 70%	0 - 36	20 - 75	LI 0.36 - 2.77
Talampanel	2010	II	59	Def, Prob	≥ 60%	0 - 24	18 - 85	-
G-CSF	2010	II	39	Def, Prob	≥ 50%	0 - 72	18 - 85	FALS
Ozanezumab	2017	II	303	All	≥ 65%	0 - 30	18 - 80	
Dexpramipexole	2013	III	942	All	≥ 65%	0 - 24	18 - 80	
Lithium	2013	III	214	All	≥ 60%	6 - 36	≥ 18	
Lithium	2010	II	84	All	≥ 60%	0 - 36	≥ 18	-
Ceftriaxone	2014	III	513	All	≥ 60%	0 - 36	≥ 18	-
Tirasemtiv	2016	II	388	All	≥ 50%	-	≥ 18	ALSFRS (2)

Exclusion rate 2010 - 2017: 0.64 (0.54-0.74)

Pooled exclusion rate: 0.60 (0.53-0.66)



Patient Registries now

- Many other crucial uses of data
  - To answer very relevant clinical questions
  - To improve development of new treatments

- EHR are transactional driven systems not designed for science: They are a source, but not the only one.

- Building into patient registries allows major leap in quality, flexibility and research focus.

- Product specific registry studies have come with many limitations.