Simulation studies in R: A case of developing novel/efficient methods for clinical trial design

Luke Ouma

Biostatistics Research Group, Faculty of Medical Sciences, Newcastle University

Oncology Biometrics, AstraZeneca []

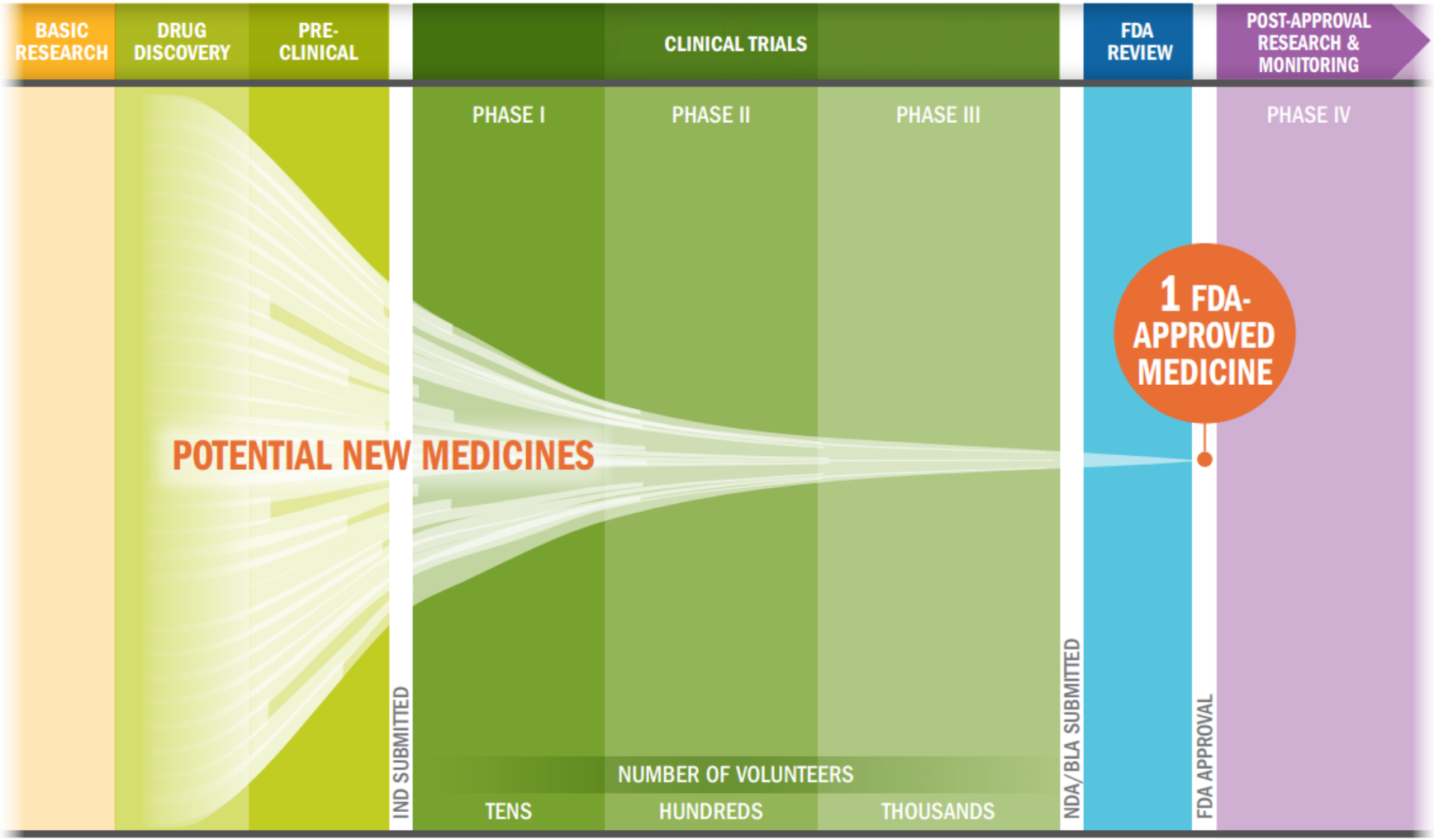
June 10, 2024

Disclosures

This presentation contains work and findings that were undertaken during my previous employment at Newcastle University. The views expressed and conclusions drawn are my own, and the content herein does not reflect or imply endorsement or affiliation with my current employer, AstraZeneca (UK).

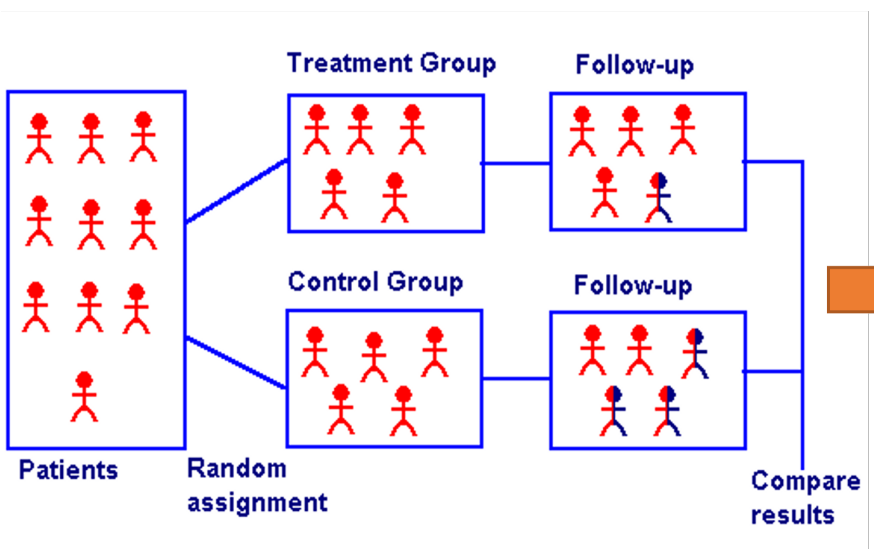
Outline

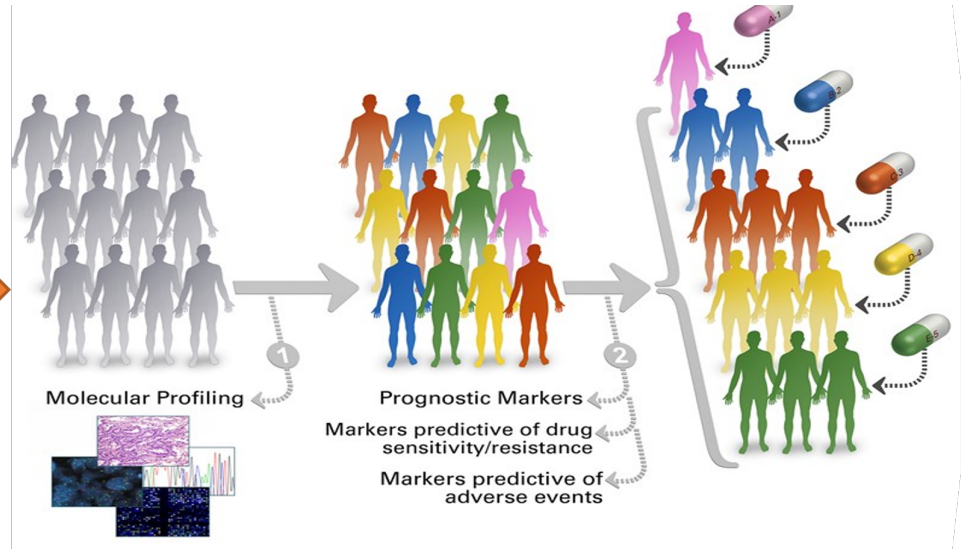
Motivation



- Q: Does a given treatment work better than control on average?
- Heterogeneity in response to treatment is fairly common → precision medicine.

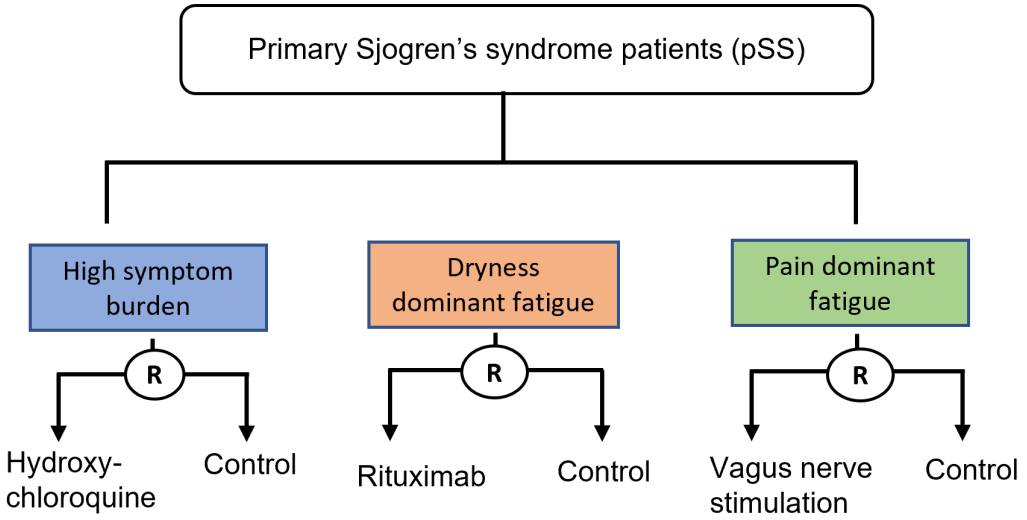
Motivation





- **Q**: Which subgroups of patients benefit from a given treatment & to what extent?
- We can investigate, multiple therapies, multiple diseases, or both under a single trial infrastructure → **Master protocols**

Motivation



- Umbrella trial - Multiple targeted therapies evaluated in a single disease setting.
- Centralised infrastructure for screening and identification of patients

Vemurafenib in Multiple Nonmelanoma Cancers with *BRAF* V600 Mutations

Table 2. Preliminary Best Response According to Cohort.*

Variable	NSCLC (N=20)	Colorectal Cancer		Cholangio-carcinoma (N=8)	ECD or LCH (N=18)	Anaplastic Thyroid Cancer (N=7)
		Vemurafenib (N=10)	Vemurafenib + Cetuximab (N= 27)			
Patients with ≥1 postbaseline assessment — no.	19	10	26	8	14	7
Complete response — no. (%)	0	0	0	0	1 (7)	1 (14)
Partial response — no. (%)	8 (42)	0	1 (4)	1 (12)	5 (36)	1 (14)
Stable disease — no. (%)	8 (42)	5 (50)	18 (69)	4 (50)	8 (57)	0
Progressive disease — no. (%)	2 (11)	5 (50)	7 (27)	3 (38)	0	4 (57)
Missing data — no. (%)†	1 (5)	0	0	0	0	1 (14)
Overall response — no. (%) [95% CI]	8 (42) [20–67]	0	1 (4) [<1–20]	1 (12) [<1–53]	6 (43) [18–71]	2 (29) [4–71]

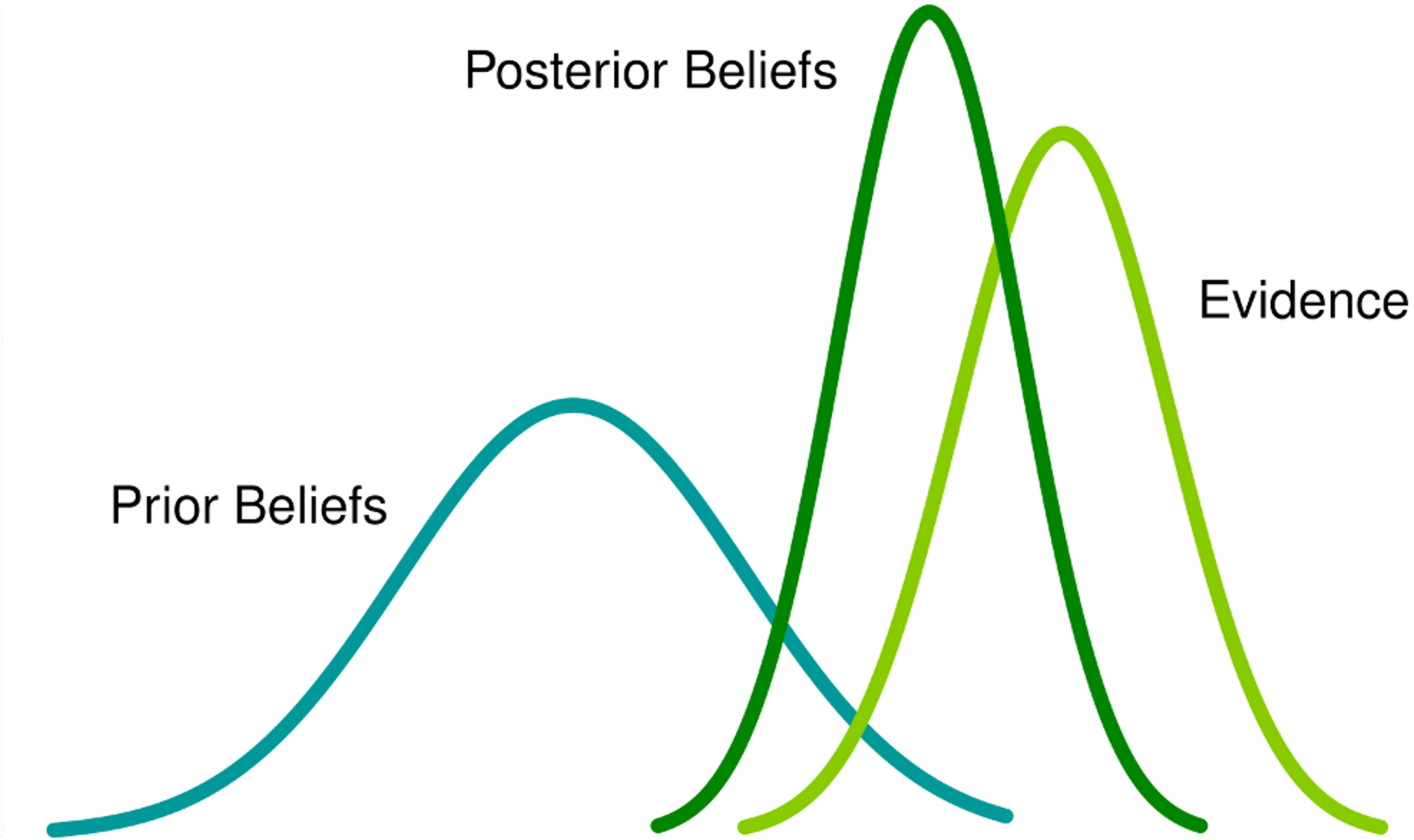
- Basket trial - Single treatment evaluated in multiple diseases that harbour a common characteristic.

Motivation

- Many master protocols test a common treatment or control.
- Rare subgroups imply randomisation is infeasible or there's need to learn more about experimental treatments.
- Practical challenges in the design such as biomarker issues.

- Umbrella and basket trials raise several additional statistical complexities in their quest to answer more therapeutic questions.
- Designing subtrials independently confers operational efficiency and logistical advantages, but misses out on potential statistical advantages.

Goal - Right treatment. Fewer patients . Less time


Posterior Beliefs
Evidence
Prior Beliefs

Simulation studies

Adapted from Morris *et al*(2019).

1. Identify *specific* **A**ims of simulation study
2. Define **D**ata-generating mechanisms
   - For instance, resampling or simulation from some parametric model.
   - For simulation from a parametric model, decide how simple or complex the model should be and whether it should be based on real data.
   - Determine what factors to vary and the levels of factors to use.
   - Decide whether factors should be varied fully factorially, partly factorially or one-at-a-time.
3. Define **E**stimands/target of analysis.
4. Carefully identify **M**ethods to be evaluated and consider whether they are appropriate for estimand/target identified.
5. List all **P**erformance measures to be estimated, justifying their relevance to estimands or other targets.
   - Give explicit formulae for the avoidance of ambiguity.
   - Choose a value of nsim that achieves acceptable Monte Carlo SE for key performance measures.

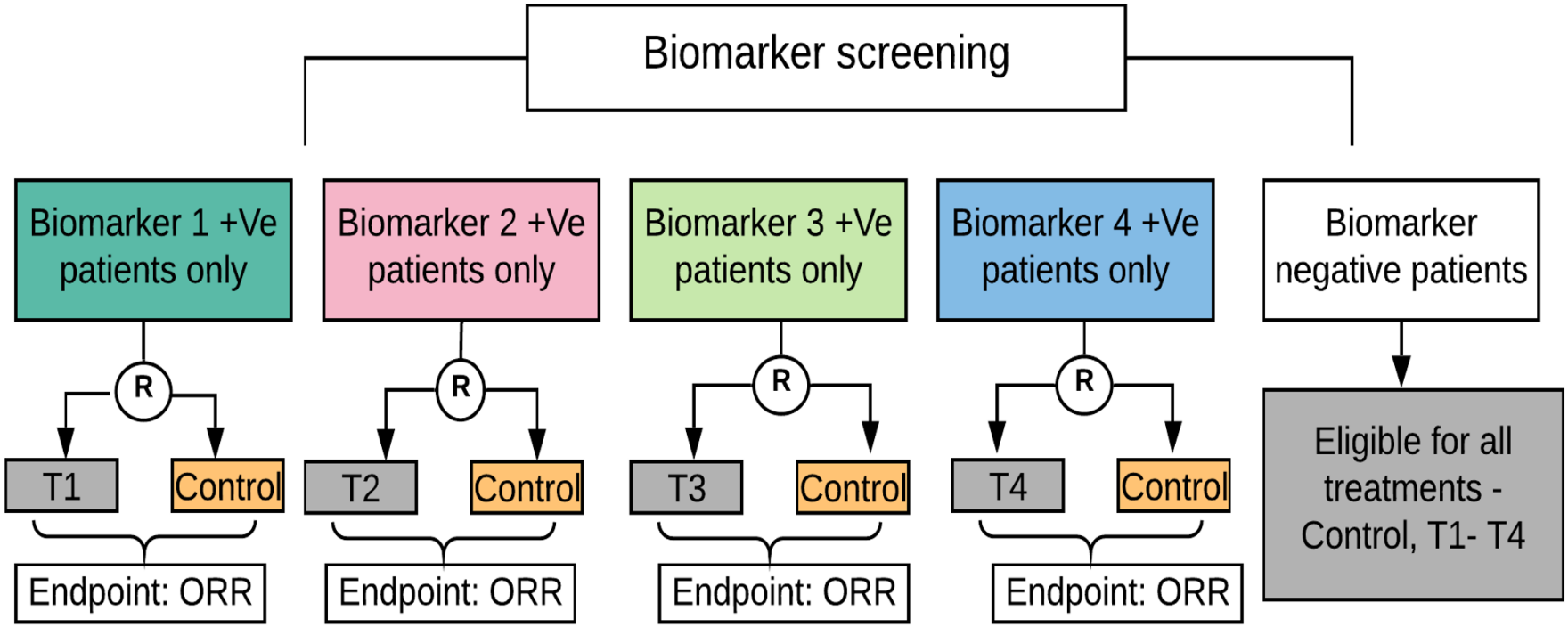
Simulation studies

- Start small and build up code, including plenty of checks.
- Set the random number seed once per simulation repetition.
- Store the random number states at the start of each repetition.
- For simulations run in parallel, use separate streams of random numbers

- Conduct exploratory analysis of results, particularly graphical exploration.
- Compute estimates of performance and Monte Carlo SEs for these estimates.

- Report simulation study using ADEMP structure.
- Graphical and tabular presentations.
- Include Monte Carlo SE as an estimate of simulation uncertainty.
- Publish code including user-written routines.

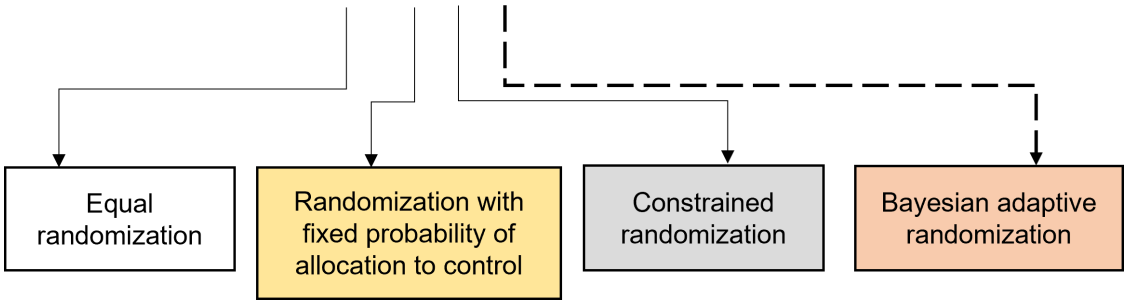
Treatment allocation in umbrella trials



- Patients often test positive for multiple biomarkers each linked to treatment.
- Sounds like a simple problem? - current practice - physician decision or randomise, sometimes hierarchy approach.
- **Q:** How do we handle eligibility to multiple subgroups? How important is treatment allocation in this context?


RANDOMIZE();


RANDOMIZE();


Equal randomization
Randomization with fixed probability of allocation to control
Constrained randomization
Bayesian adaptive randomization

- Trial design : Binary endpoint, 4 biomarkers, each with linked treatment, $T_1 - T_4$; Patients with only one marker eligible for control (T_0) and linked treatment $T_j, j = 1, ...4$; marker prevalence - $B_1, B_3 = 0.3$; $B_2, B_4 = 0.25$
- n_{sc} different scenarios; 10,000 simulation replicates.
- Evaluate the bias & MSE; statistical power; the average number and proportion of (i) patients on an experimental treatment; (ii) patients on the best treatment available to them; (iii) patient responses; and (iv) patients on each treatment.

- https://github.com/oondijo/multipleBiomarkers

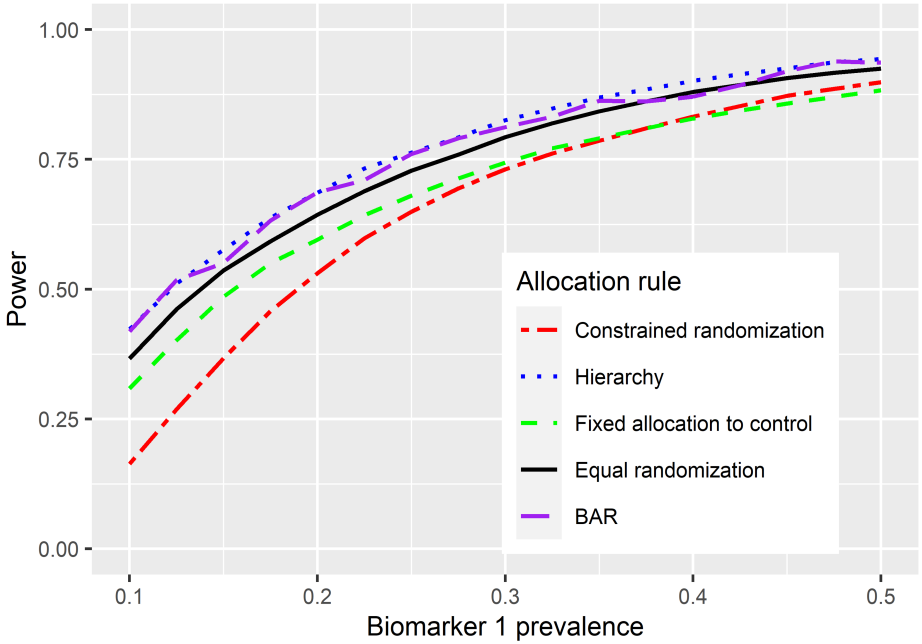


Figure: Statistical power of the five-treatment allocation approaches as B_1 prevalence varies

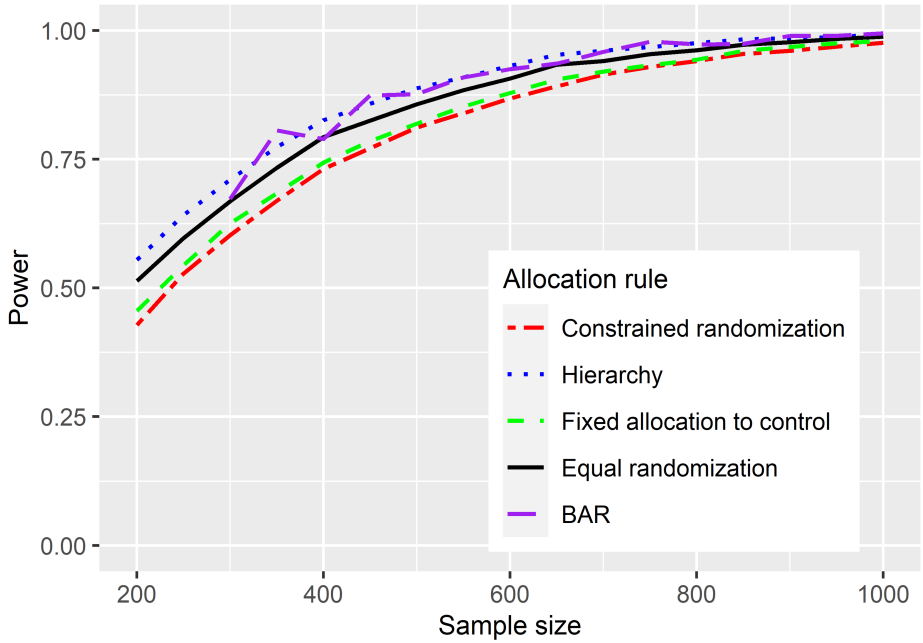


Figure: Statistical power of the five-treatment allocation approaches as sample size varies

Table: A comparison of the average proportion of patients on experimental treatment for the five treatment allocation strategies across all scenarios

Treatment allocation strategy		**Proportion (Range*)**
Equal Randomisation		69.3% (52.7–72.5)
RFAC	$\vartheta = 0.2$	80.0% (71.2–87.5)
	$\vartheta = 0.25$	75.0% (66.0–83.0)
	$\vartheta = 0.3$	70.0% (60.5–78.8)
Hierarchy	$\rho = 0.5$	64.5% (54.8–72.7)
	$\rho = 0.75$	57.3% (47.0–66.3)
	$\rho = 0.9$	52.9% (43.7–62.5)
CR	$\phi = 0.5$	79.0% (72.0–80.5)
	$\phi = 0.75$	79.8% (78.8–80.2)
	$\phi = 0.9$	79.9% (79.3–80.2)
BAR		62.6% (54.0-70.3)

Note: RFAC: Randomisation with a fixed allocation probability to control; CR: Constrained randomisation; BAR: Bayesian adaptive randomisation. *The range is across the 12 different simulation scenarios

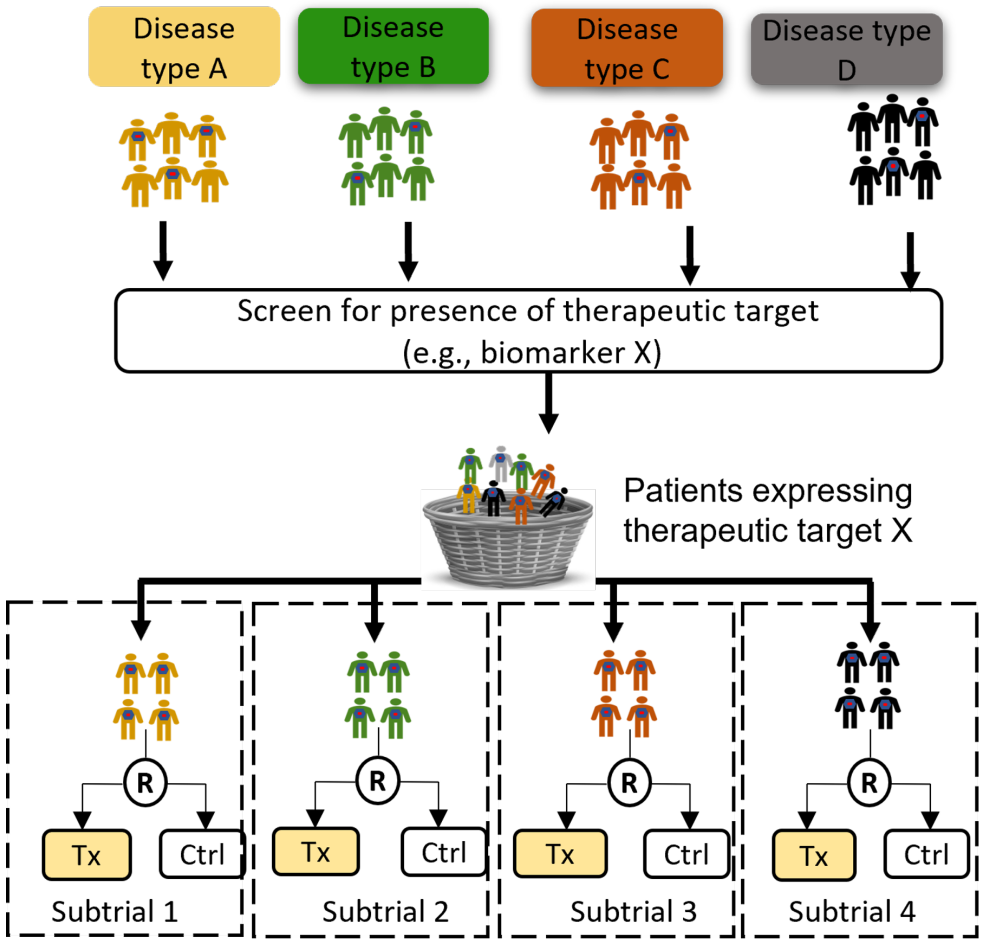
Hierarchy of biomarkers
Highest power
Dependent on degree of belief on the validity of hierarchy

CR approach
Balancing allocation to treatment arms when recruitment is low to some arms

BAR
- Highest power
- Highest proportion on the best treatment available to them if a treatment delivers an unanticipated detrimental effect

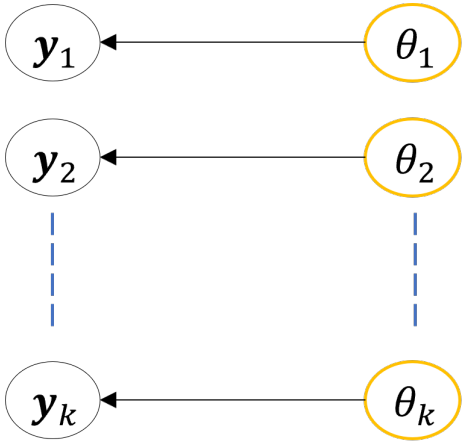
- Pre-specification of a treatment allocation approach in umbrella trials is necessary.
- Trial sample size, biomarker prevalence, and prevalence of individual overlaps within the patient population are significant considerations in choosing an approach.

Analysis of randomised basket trials



Randomised basket trial design

- **Q**: What is the best analysis strategy for randomised basket trials?



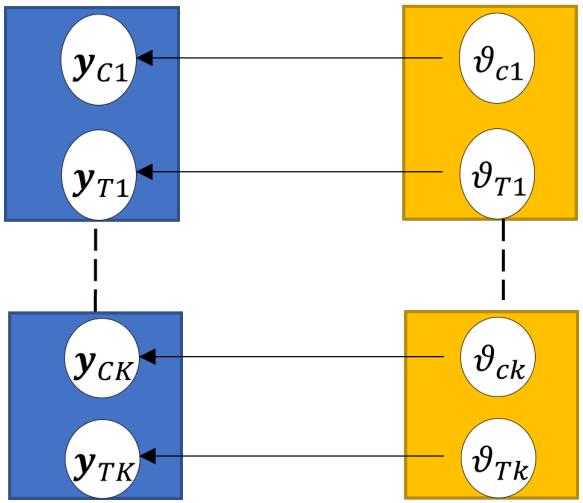
- Standalone analyses & complete pooling —widely criticised
- Robust borrowing of information—commensurate prior approach.

$$E(y_{ij}) = f(X_{ij}, \theta_i), j = 1, ..., n_i$$

$$\theta_i|\theta_q, \nu_{iq} \sim N(\theta_q, \nu_{iq}^{-1}) \ \forall \ i = 1, ..., k)$$

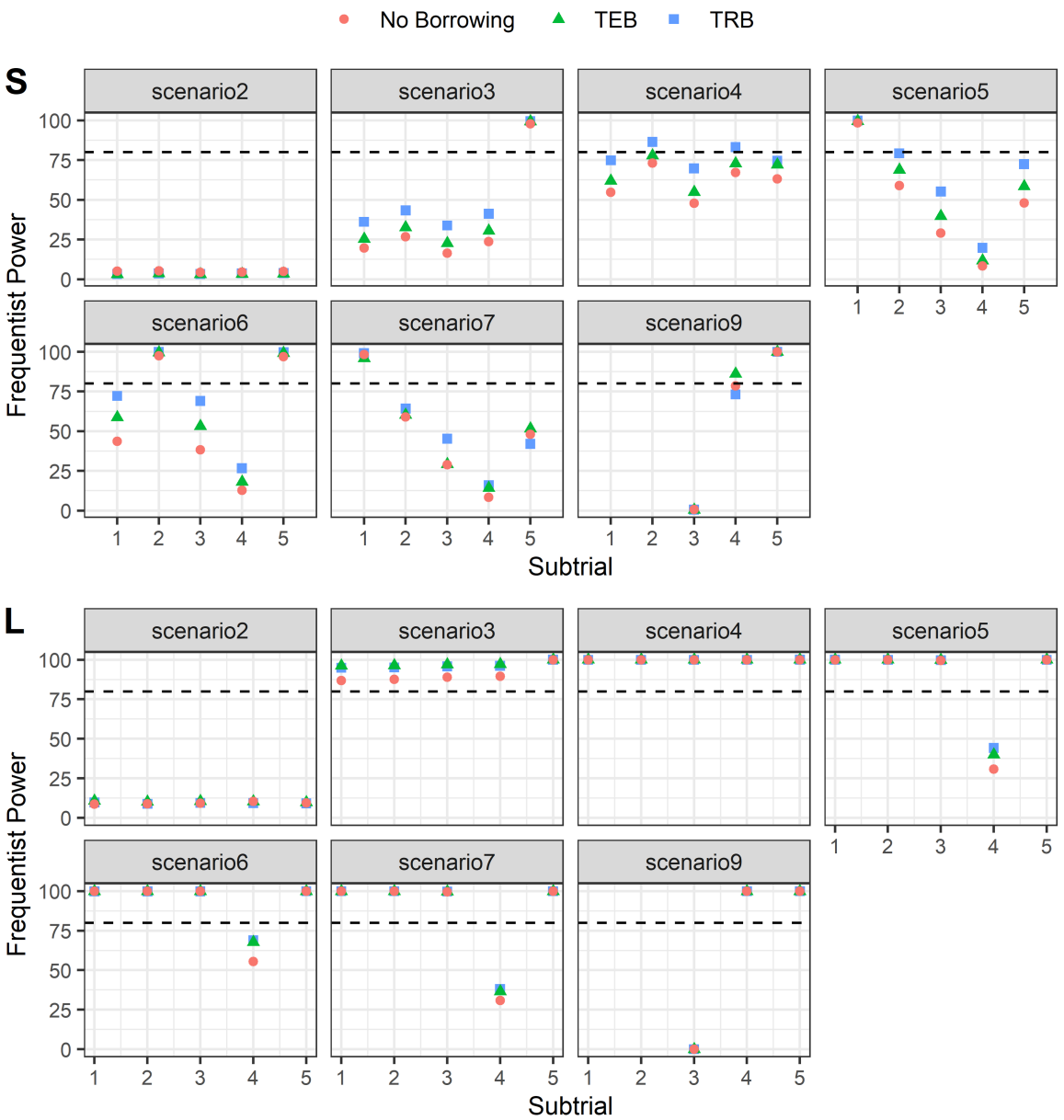
$$\nu_{iq} \sim \omega_{iq}\pi_1(\nu_{iq}) + (1 - \omega_{iq})\pi_2(\nu_{iq}), \text{ with } q \neq i$$

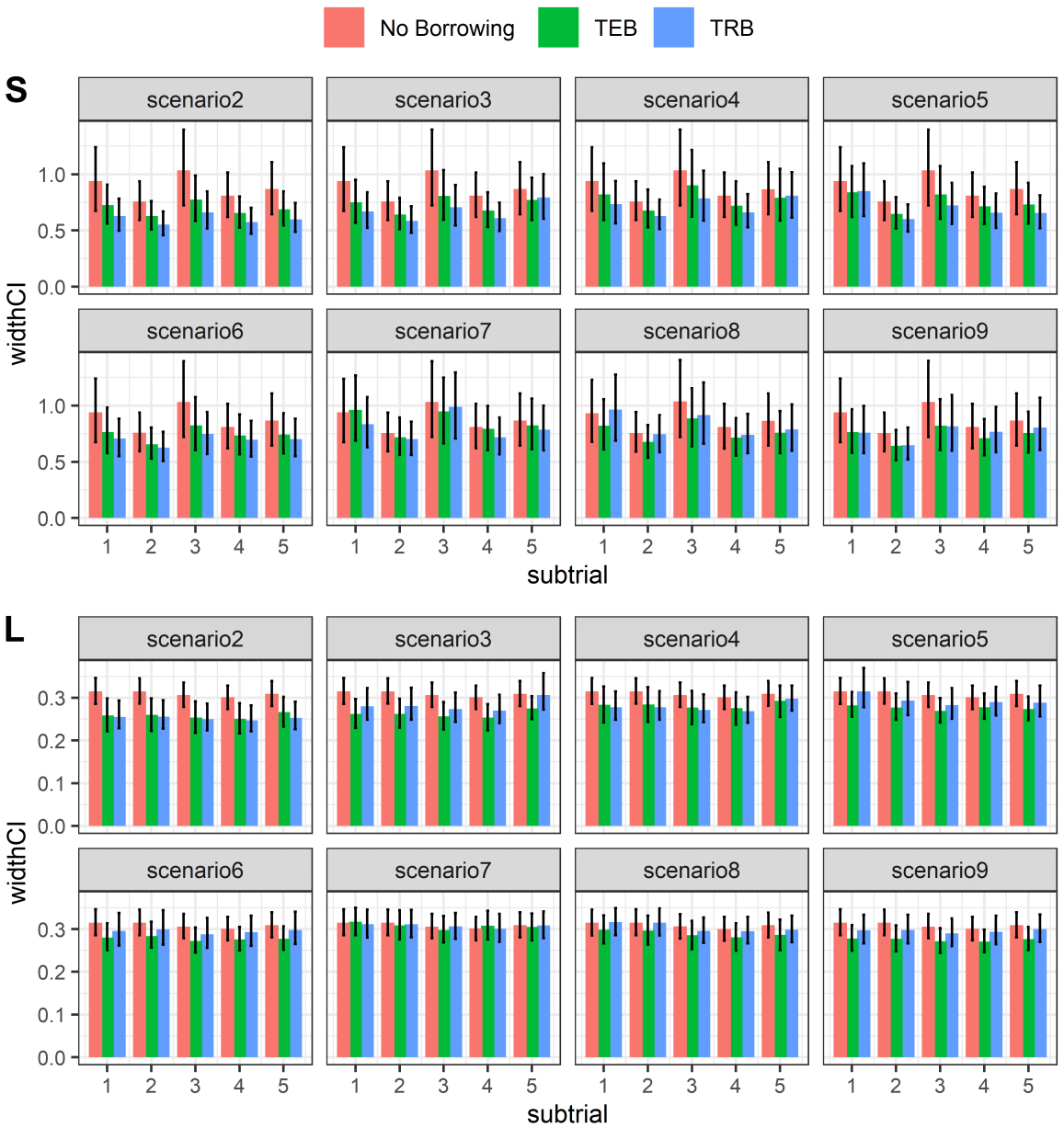
- Treatment effect borrowing (TEB)[1] and Treatment response borrowing (TRB) alternatives in RBT.



- Using a distributional discrepancy, we characterise the pairwise commensurability between θ_i and θ_q.
- Obtain a (collective) predictive prior $\pi_0(\theta_i|y_{(-i)}) \rightarrow \pi_p(\theta_i|y_{(-i)}, y_i)$ and then characterise the relative importance of each of the $k-1$ complementary subtrials.

- Basic trial settings: Phase II basket trial with $K = 5$ subtrials. Maximum of 336 patients recruited, $n_1 = 70, n_2 = 66, n_3 = 64, n_4 = n_5 = 68$.
- Simulate data as follows; we suppose that $y_{ik} \sim N(\mu_{ik}, \sigma^2)$ with $\mu_{ik} = \beta_{0k} + \theta_k T_{ik}$. We assume $\beta_{0k} = 5$ and the inter-patient standard deviation $\sigma = 0.4$.
- 9 different scenarios evaluated.


No Borrowing
TEB
TRB
S
scenario2
scenario3
scenario4
scenario5
scenario6
scenario7
scenario9
Frequentist Power
Subtrial
L
scenario2
scenario3
scenario4
scenario5
scenario6
scenario7
scenario9
Frequentist Power
Subtrial


No Borrowing
TEB
TRB
S
L
scenario2
scenario3
scenario4
scenario5
scenario6
scenario7
scenario8
scenario9
widthCI
subtrial

- **Case study 1:** Diet and exercise trial – 60 participants randomised to 3 exercise regimes and within each regime further randomised to diet1 and diet 2.
- **Case study 2:** A multicentre, randomised, double-blind, placebo-controlled, proof-of-concept study of iscalimab in patients with primary Sjögren's syndrome (Fisher et al., 2020).

Table: Posterior probability that θ_k exceeds a pre-specified threshold, $P(\theta_k > \delta | data)$

Case study	**Modelling strategy**	**S1**	**S2**	**S3**
Case study 1 All $n_k = 35$	NB	0.4185	0.6145	
	TRB	0.2750	0.8515	1.000
	TEB	0.2390	0.3975	1.000
Case study 2 All $n_k = 20$	NB	0.016	0.2465	-
	TRB	0.039	0.365	-
	TEB	0.0275	0.468	-

Case study	**Modelling strategy**	**Subtrial**		
		S1	S2	S3
Case study 1 All $n_k = 35$	NB	2.50 (-2.11, 7.07)	4.0 (-2.83, 10.75)	23.59 (16.39, 30.01)
	TRB	1.73 (-2.81, 5.97)	6.27 (-0.29, 12.30)	27.45 (19.96, 34.98)
	TEB	1.31 (-0.12, 5.46)	2.23 (-3.63, 7.85)	33.27 (23.91, 48.71)
Case study 2 All $n_k = 20$	NB	0.21 (-2.26, 2.72)	-3.10 (-4.65, -1.46)	-
	TRB	-0.52 (-2.60, 1.55)	-2.73 (-4.18, -1.19)	-
	TEB	-0.46 (-2.61, 1.74)	-2.59 (-4.21, -0.82)	-

Table: Posterior mean treatment effects, θ_k (95% credible interval)

- Proposed a new strategy for analysis of randomised basket trials.
- Both modelling strategies provide substantial efficiency gains – higher power and precision of effect estimates.
- TRB outperforms TEB especially when subtrial sample sizes are small on all operational characteristics.
- TEB has considerable gains in performance over TRB when subtrial sample sizes are large, or the treatment effects and groupwise mean responses are noticeably heterogeneous across subtrials.
- TRB, and TEB can potentially lead to different conclusions in the analysis of real data.
- Borrowing should be carefully considered/implemented depending on the potential heterogeneity in effects across subgroups and the subtrial sample sizes.

Conclusion

- Why the code/script/program and its intended purpose.
- Limit hard coding - defining variables, input and output files at the onset.
- START SMALL! - Segregating distinct code
- Collaborate - practice code review.

Conclusion

- Why the code/script/program and its intended purpose.
- Limit hard coding - defining variables, input and output files at the onset.
- START SMALL! - Segregating distinct code
- Collaborate - practice code review.
- Elegant code/programs vs time - develop what works before you automate, otherwise you lose $n + 1$hrs on each project.

Conclusion

- Precision medicine trials offer considerable efficiencies in drug development, but practical challenges in the design and analysis raise complexities that necessitate new statistical methodology.

Better methods, better trials.

frontiers | Frontiers in Medicine

TYPE Systematic Review
PUBLISHED 12 October 2022
DOI 10.3389/fmed.2022.1037439

Check for updates

OPEN ACCESS

EDITED BY
Cristiana Sessa,
Oncology Institute of Southern Switzerland (IOSI), Switzerland

REVIEWED BY
Valter Torri,
Mario Negri Pharmacological Research Institute (IRCCS), Italy
Hideki Maeda,
Meiji Pharmaceutical University, Japan

Design and analysis of umbrella trials: Where do we stand?

Luke O. Ouma[1]*, James M. S. Wason[1], Haiyan Zheng[2], Nina Wilson[1] and Michael Grayling[1]

[1]Biostatistics Research Group, Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom, [2]Medical Research Council (MRC) Biostatistics Unit, University of Cambridge, Cambridge, United Kingdom

Received: 27 January 2022 | Accepted: 1 September 2022

DOI: 10.1111/rssc.12602

ORIGINAL ARTICLE

Royal Statistical Society Series C Applied Statistics

Bayesian modelling strategies for borrowing of information in randomised basket trials

Luke O. Ouma[1] | Michael J. Grayling[1] | James M. S. Wason[1] | Haiyan Zheng[2]

Received: 28 August 2020 | Revised: 1 February 2021 | Accepted: 13 March 2021

DOI: 10.1002/pst.2119

MAIN PAPER

WILEY

Treatment allocation strategies for umbrella trials in the presence of multiple biomarkers: A comparison of methods

Luke Ondijo Ouma[1] | Michael J. Grayling[1] | Haiyan Zheng[1] | James Wason[1,2]

Acknowledgements

Newcastle University

Faculty of Medical Sciences

List of References

[1] Haiyan Zheng and James MS Wason. "Borrowing of information across patient subgroups in a basket trial based on distributional discrepancy". In: *Biostatistics* 23.1 (2022), pp. 120–135.

Thank You