Innovative trial designs

Glossary for common terms and trial designs

Australian Clinical Trials Alliance

Innovative Trial Design Working Group

Version 0.0.1

Preface

This glossary contains definitions for terms and trial designs that are commonly encountered in the broadly defined field of 'innovative trial designs'. The glossary is intended to fulfil the objectives of the nomenclature component of the Australian Clinical Trials Alliance (ACTA) Innovative Trial Design Working Group. Readers who may find the glossary useful include other ACTA working groups, human research and ethics committees, consumers and community, and new entrants into the field of innovative trial design including students and clinical researchers. The intention is for this glossary to be a centralised, living document that will be updated and improved upon either suggestion or indication, with the general consensus of the ACTA Innovative Trial Design Working Group, as the field of innovative trial design grows and matures.

For the purposes of this glossary, innovative trial designs include adaptive trials that focus on either between-patient randomisation (such as multi-arm, multi-stage designs or response adaptive randomisation) or within-patient randomisation (such as sequential multiple assignment randomised trial designs), platform trials (including basket and umbrella trials), cluster randomised trials (including stepped-wedge and cross-over trials), and embedded designs (including registry trials).

Comments, corrections, and suggested additions can be emailed directly to the ACTA Innovative Trial Design Working Group via: acta@clinicaltrialsalliance.org.au

Contents

A **6**
Adaptive . . . 6
Adaptive treatment policy . . . 6
Alpha spending . . . 6

B **6**
Basket trial . . . 6
Bayesian inference . . . 6
Borrowing . . . 6

C **6**
Concurrently randomised cohort (CRC) . . . 6
Consumer . . . 6
Contemporaneous cohort . . . 6
Covariate . . . 6
Covariate-adjusted adaptive randomisation . . . 7
Cluster randomisation . . . 7
Confidence interval . . . 7
Core protocol . . . 7
Credible interval . . . 7

D **7**
Data safety and monitoring board (DSMB) . . . 7
Decision quantity . . . 7
Decision rule . . . 7
Decision rule threshold . . . 7
Decision-theoretic . . . 7
Domain . . . 8
Domain-specific appendix . . . 8
Dynamic treatment regimen . . . 8

E **8**
Early stopping . . . 8
Embedding . . . 8
Embedded fixed design . . . 8
Estimand . . . 8
Estimand framework . . . 8
Enrichment . . . 8
Epoch . . . 8

F **8**
Futility . . . 8
Frequentist inference . . . 8

G **8**
Group sequential . . . 8

H **8**
Hierarchical model . . . 8
Historical control . . . 9

I **9**
Inferiority . . . 9
Interim analysis . . . 9

J **9**

K **9**

L **9**
Likelihood function . . . 9
Likelihood principle . . . 9
Loss function . . . 9

M **9**
Master protocol . . . 9
Markov-chain Monte Carlo (MCMC) . . . 9
Minimisation . . . 9
Monte Carlo . . . 9
Multi-arm, multi-stage (MAMS) . . . 9
Multiplicity . . . 9
Multiple imputation . . . 9

N **9**
Network meta-analysis . . . 9
Non-concurrent control . . . 9
Non-inferiority . . . 10
Null hypothesis . . . 10
Null hypothesis significance testing (NHST) . . . 10
N-of-1 . . . 10

O **10**
Operating characteristics . . . 10
Outcome adaptive randomisation . . . 10

P **10**
Participant randomisation scheme . . . 10
Participant reported outcome . . . 10
Platform trial . . . 10
Posterior distribution . . . 10
Posterior sampling . . . 10
Prior distribution . . . 10
Posterior predictive distribution . . . 10
Population . . . 10
Population summary . . . 10
p-value . . . 10

Q **10**

R **10**
Real-world data . . . 10
Real-world evidence . . . 11
Regimen . . . 11
Registry . . . 11
Registry trial . . . 11
Randomisation . . . 11
Response adaptive randomisation . . . 11
Between-participant . . . 11
Within-participant . . . 11

S **11**
Sample size re-estimation . . . 11
Scheduled analysis . . . 11
Sequential multiple assignment randomised trial (SMART) . . . 11
Simulation . . . 11
Silo . . . 11
Stage . . . 11

State . . . 11
Statistical analysis plan (SAP) . . . 11
Statistical appendix . . . 11
Statistical significance . . . 12
Strata . . . 12
Subgroup . . . 12
Superiority . . . 12
Stepped wedge . . . 12

T **12**
Terminal analysis . . . 12
Thompson sampling . . . 12
Traditional trial . . . 12
Trial steering committee . . . 12

U **12**
Util . . . 12
Utility . . . 12
Utility function . . . 12
Umbrella trial . . . 12
Unit of analysis . . . 12

V **12**
Version control . . . 12
Value-of-information . . . 12

W **12**

X **13**

Y **13**

Z **13**

A

Adaptive As in 'adaptive trial' or 'adaptive design'. An adaptive trial is a trial where aspects of the trial design may be modified, typically based on analysis of accumulated data at interim analyses. Can be either Bayesian or frequentist. See interim analysis.

General overviews of adaptive trial designs, including common design types, can be found in Berry (2012), Park et al. (2018), and Pallmann et al. (2018). For a clinician-focused primer on design consideration, see Thorland et al. (2018). For an overview of methodological considerations, see Granholm et al. (2022). Berry et al. (2010) provide a comprehensive handbook on adaptive trials using Bayesian methods in particular. For a regulators perspective, see US Food and Drug Administration (2010, 2018, 2020, 2022). Wason et al. (2019) provide an insightful commentary on when adaptive designs are *not* useful.

Adaptive treatment policy See dynamic treatment regimen.

Alpha spending A flexible, frequentist approach of distributing (or spending) the type I error (or alpha) over the duration of a group sequential trial. It allows for new looks to be added without inflating the overall level of type I error.

B

Basket trial The US Food and Drug Administration (2022) define a basket trial as a:

[...] trial [that] involves a single investigational drug or drug combination that is studied across multiple populations defined by disease, histology, number of prior therapies, genetic or other biomarkers, or demographic characteristics.

Note that, more generally, the intervention does not necessarily need to be a drug and could conceivably be any type of intervention.

Bayesian inference A gentle definition of Bayesian inference is given by Gelman et al. (2014):

Bayesian inference is the process of fitting a probability model to a set of data and summarising the result by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations.

A more technical definition is that it is an approach to statistical inference that conditions on observed data using a likelihood function and treats all other quantities, including parameters, as random variables. A necessary condition for Bayesian inference is that parameters of the associated likelihood function are assigned prior distributions. All inference is based on the posterior distribution, which combines the information from the likelihood and the prior distributions. Generally speaking, the alternative to Bayesian inference is frequentist inference.

Borrowing An informal shorthand for a Bayesian technique that uses relevant information from other subgroups within a trial, typically via a hierarchical model, to provide a more informative estimate of efficacy in other trial subgroups of interest.

McGlothlin and Viele (2018) provide a non-technical description borrowing using Bayesian hierarchical models. Berry et al. (2013) describe the technique in the context of phase II oncology trials and Murthy et al. (2021) make the case for borrowing between paediatric and adult cohorts within a single clinical trial.

C

Concurrently randomised cohort (CRC) Trial participants whose members had the same treatments available to them and the same chance of receiving those treatments at the same time. A fixed trial design has a single concurrently randomised cohort. In an adaptive trial, if the set of interventions that a patient can be randomised to changes, there will be multiple contiguous concurrently randomised cohorts. See also contemporaneous cohort, non-concurrent controls, and stage.

Bofill-Roig et al. (2022), Saville et al. (2022), Marschner et al. (2022), and Wason et al. (2022) provide different strategies for analysis.

Consumer Patients and potential patients, carers, and people who use health care services. Consumers can also be people who represent the views and interests of a consumer organisation, a community, or a wider constituency.

Contemporaneous cohort Trial participants who were randomised within the same stage of an adaptive trial. This is different to concurrently randomised cohort. Participants may be contemporaneous but not concurrent, and vice versa. See concurrently randomised cohort, non-concurrent controls, and embedded fixed design.

Covariate Can be either baseline (pre-randomisation) or post-randomisation patient

characteristics distinct from the interventions.

Covariate-adjusted adaptive randomisation A method of randomisation aiming to achieve balanced treatment allocation over a set of influential covariates.

Cluster randomisation When treatment is randomly allocated to groups (or clusters), rather than individuals. Used as the basis for cluster randomised trials including stepped-wedge and cluster cross-over designs. In a cluster randomised trial, the 'unit of analysis' typically refers to the level at which the data are analysed. For example, the 'unit of analysis' may be at the individual level, but in some cases outcomes may be analysed at the cluster level.

Confidence interval A probabilistic interval that covers the true value of the population parameter of interest with a nominal degree of confidence, e.g. 95%. Confidence intervals based on repeatedly samples of the population will contain the true value 95% of the time.

Core protocol Also known as a 'master protocol'. A document that details the central aims of a trial along with core trial endpoints, decision rules, estimands, and trial governance structures that will be consistent throughout the design and across different domains. Typically used in platform, umbrella, and basket trials. Ideally immutable over time, although subject to change with appropriate approvals from ethical, research, and funding bodies. The core protocol is typically supplemented by appendices that describe specific aspects of the trial, facilitate design adaptations, typically including a statistical appendix, a continually updated and version controlled implementation guide for scheduled analyses, and statistical analysis plans that are implemented when a terminal analysis is required. See recent regulatory guidelines (US Food and Drug Administration, 2020).

Credible interval The interval within which the value of a population parameter of interest belongs with some probability. Used to summarise the uncertainty around a parameter. Typically, an equal-tailed 95% credible interval is used for inference, which is the interval in which the probability in each tail outside of the interval is 0.025. Often abbreviated to CrI to distinguish from frequentist confidence intervals. See 'Bayesian inference'.

D

Data safety and monitoring board (DSMB) Also known as a data safety and monitoring committee (DSMC). The DSMB is a committee independent of both the Trial Steering Committee and any trial sponsor or funder. It advises these bodies on continuation or stopping based upon safety and efficacy considerations. The primary objective of the DSMB is to assure safety for the patients in the trial. Regulatory guidance has been long established by the US Food and Drug Administration (2006). For an insightful historical note on the evolution of DSMBs, see Meinert (2022)

Decision quantity A measure of evidence that is used to make trial decisions, for example arm stopping in an adaptive trial. This quantity can be either Bayesian (e.g. the probability of an inequality between population summary between two intervention groups with a scalar value) or frequentist (e.g. a p-value).

Decision rule A statistical rule that states the point at which a decision quantity passes a decision rule threshold, at which point following the rule must be recommended to the DSMB. Typically, decision rules are constructed to represent the either superiority, equivalence, inferiority, or the non-inferiority of an intervention, or the statistical futility of continuing the trial to reach any of these conclusions. Can be either Bayesian or frequentist.

Decision rule threshold The specific threshold, applied to a quantity of interest, at which a decision rule is to be followed. For example, a nominal 95% threshold might be used to construct a Bayesian decision rule where a posterior probability of a treatment effect being positive is greater than 0.95, or frequentist decision rule a p-value less than 0.05.

Decision-theoretic Refers to the use of an explicit utility (or loss) function that incorporates the value of outcomes with their probability to make decisions under uncertainty. Can be either Bayesian or frequentist. See loss function, utility function, and value-of-information.

Seminal works in decision-theory extend back to Von Neumann and Morgenstern (1944) who developed the general theory, and Savage (1954) who formalised the statistical theory. Berger (1980) and Robert (2007) provide comprehensive technical references for statistical decision-theoretic methods with a mostly Bayesian focus. Relevant to clinical trial design, Claxton et al. (2000) provide an accessible formalisation of treatment decisions and trial design, Lipsky and Lewis (2013) describe design and ethics of decision-theoretic response adaptive randomisation, Ryan et al. (2016) provide a review of modern computation algorithms for Bayesian

decision-theoretic experimental design, and Hee et al. (2016) review decision-theoretic designs of small trials and pilot studies.

Domain A set of mutually exclusive and competing interventions that share a common clinical mode of action or clinical context of use.

Domain-specific appendix An appendix to a core or master-protocol document describing the protocol relating to a given domain of a multifactorial platform trial.

Dynamic treatment regimen A set of rules describing how treatment could be assigned in response to some dynamically changing factor, for example, degree of treatment response. Identifying the optimal dynamic treatment regimen from those under consideration is desired. Sequential multiple assignment randomised trials can be used to obtain data to identify optimal dynamic treatment regimens. Also known as a dynamic treatment regime or an adaptive treatment policy. Arguably first described by Robins in 1986, with more recent pioneering contributions around the time of Lavori and Dawson (2000) and Murphy (2003). Several excellent texts are available detailing the appropriate statistical methodology for dynamic treatment regimens including Chakraborty and Moodie (2013 and Tsiatis et al. (2019).

E

Early stopping The process of stopping a trial, or a component of a trial, on the basis of accumulated data, before the trial was planned to terminate. See adaptive design, response adaptive randomisation, and multi-arm multi-stage designs.

Embedding The process of integrating research activities within routine patient care, e.g one or more of screening, recruitment, delivery of intervention, and data collection.

Embedded fixed design One of the unique randomisation configurations occuring within a participant randomisation scheme, which is equivalent to a fixed design embedded within an overall adaptive trial.

Estimand Specific definition of the quantity of interest in a research study. Defining an estimand entails providing details on five attributes: intervention condition(s), population, outcome, population-level summary and post-randomisation events (also known as intercurrent events).

Estimand framework A systematic approach to thinking through the trial objectives to ensure that the trial goals are both precise and transparent (through the specification of the estimand) and that the proposed design and analysis are aligned with them.

Enrichment A study design that allow researchers to identify sub-populations of participants for whom a proposed intervention is more likely to be beneficial and increase enrolment from those sub-populations.

Epoch An element of a set of mutually exclusive and exhaustive time periods that is typically used in statistical modelling for clinical trials with response adaptive randomisation or early-stopping to either incorporate non-concurrent control groups or treatment dynamics (with respect to time), or as a complement to a piecewise analysis of concurrently randomised cohorts. See also concurrently randomised cohorts, contemporaneous cohorts, and non-concurrent controls.

F

Futility A statistical decision rule at which point the there is little chance or no chance of demonstrating a clinically meaningful effect (e.g. superiority) if the trial was to continue as planned.

Frequentist inference Commonplace approach to statistics developed around the classical concepts of long-run probability. Forms the basis of the widely used null hypothesis significance testing framework. See also Bayesian inference.

G

Group sequential A particular type of trial design where data are examined at interim analyses and decisions made to stop or continue the intervention under investigation without inflating the type 1 error. See also early stopping.

H

Hierarchical model A statistical model that allows for clustered data (e.g. repeated measures on individuals, or data for individuals that belong to a group). Also commonly described as 'mixed effects', 'random effects', 'variance components', 'varying-intercept', or 'varying-slope' models. Can be either Bayesian or frequentist. For Bayesian approaches to hierachical models see Gelman et al (2014) for general modelling

strategies, and McGlothlin and Viele (2018 for more of an adaptive trial context.

Historical control Refers to data from external and often completed studies. Often used to estimate indirect treatment comparisons using trial data.

I

Inferiority A conclusion that an intervention leads to a worse outcome than some reference treatment.

Interim analysis An analysis conducted over the course of an ongoing trial using the data accumulated so far. Results can be used to implement *a priori* decision rules and potentially modify the conduct of the trial. See also 'scheduled analysis'.

J

Intentionally blank

K

Intentionally blank

L

Likelihood function Describes the joint probability of the observed data as a function of the parameters of the chosen statistical model. Maximum likelihood estimation is the cornerstone of frequentist methods.

Likelihood principle A philosophical principle stating that the only data that should affect decision making is that which is observed. Fundamental to Bayesian decision making. For more technical discussion of the likelihood principle, refer to Berger (1980) and Robert (2007).

Loss function The inverse of a utility function. See utility function.

M

Master protocol See core protocol.

Markov-chain Monte Carlo (MCMC) A simulation technique used in Bayesian statistics to approximate the posterior distribution of a parameter of interest and to sample from it. For a relatively gentle introduction to modern MCMC methods, see Gelman et al. (2014).

Minimisation A particular method of covariate-adaptive randomisation that allocates participants to the intervention group that best maintains balance in given prognostic factors, typically with some uncertainty maintained. First proposed by Pocock and Simon (1975).

Monte Carlo Simulation technique using random sampling and statistical modelling to evaluate the performance of a design; or more generally a procedure, and study its performance.

Multi-arm, multi-stage (MAMS) A common name for an adaptive trial with more than 2 arms where data are used to evaluate pre-specified decision rules at different successive interim analyses. See early stopping and response adaptive randomisation.

Multiplicity The potential inflation of the type 1 error rate as a result of multiple testing due, for instance, to comparisons across multiple arms or repeated testing of the same outcome at different times over the course of a trial.

Many methods are available to control for multiplicity in clinical trials, however opinions about the need and methods to differ among practitioners (Pike et al., 2022). For a non-technical overview, see Bender and Lange (2001). For an overview of modern approaches, including graphical methods, see Wang et al. (2015). Ryan et al. (2020) provide recommendations as to whether Bayesian adaptive designs should adjust for multiple interim analyses. Guidelines from a regulators perspective are available from the US Food and Drug Administation (2022).

Multiple imputation A statistical approach used to handle missing data. It aims to allow for the uncertainty about the missing data by creating several completed datasets where the missing values are imputed based on the observed data and appropriately combining parameter estimates obtained for each of them.

N

Network meta-analysis A type of meta-analysis where 3 or more treatments are being compared using both direct comparisons, possibly from randomised controlled trials, and indirect comparisons from historical controls in trials based on the same or similar comparator.

Non-concurrent control Data from a control arm of an adaptive trial that is not part of a concurrently randomised cohort from which an effect estimate is sought. Note that a non-concurrent control can be contemporaneous.

Can improve trial efficiency at the cost of increased bias because of time trends. See concurrently randomised cohort, and contemporaneous cohort.

Non-inferiority A conclusion that an intervention does not lead to worse outcomes than a reference treatment up to a certain limit, called non-inferiority margin. A one-sided version of an equivalence trial.

Null hypothesis An *a priori* assumption of no effect between two treatment arms. For example, typical for a continuous endpoint, a null hypothesis is assumed to be that in which a mean difference between two treatment arms is equal to exactly zero (with a fixed population variance).

Null hypothesis significance testing (NHST) A statistical hypothesis testing framework that uses frequentist methods of statistical inference used to decide whether the sampled data support a particular hypothesis. Null hypothesis significance testing assumes that the null hypothesis is true (e.g. a treatment under investigation has no effect on the outcome of interest in a clinical trial).

N-of-1 Single participant trials whereby the participant randomly receives each treatment in a randomised sequence, with participants acting as their own control.

O

Operating characteristics The long-run properties of a clinical trial design, typically power, under different scenarios. See also simulation.

Outcome adaptive randomisation See response adaptive randomisation.

P

Participant randomisation scheme An individual listing that captures, for each participant, the treatments that were available to them and their chance of receiving those treatments.

Participant reported outcome Any report of the status of a participant's health condition that comes directly from the participant, without interpretation of the participant's response by a clinician or anyone else.

Platform trial A trial design that allows multiple interventions within one or more domains across one or more subgroups of participants, that is governed centrally using a core protocol.

Posterior distribution The posterior distribution (or simply posterior) is used in Bayesian analysis to describe the information about a parameter of interest (e.g. intervention effect) after observing data. It is typically a combination of the prior distribution and new evidence provided by the likelihood function. It can be calculated at an interim analysis or any time an update is desirable.

Posterior sampling See Thompson sampling.

Prior distribution Also referred to commonly as a 'prior'. Represents the investigator's belief about the true value of a parameter, without knowledge of the observed sample data.

Posterior predictive distribution The distribution of unobserved data based on observed data and a Bayesian model (including a prior distribution). Often used in clinical trials to predict the results at the end of the trial, based on the posterior distribution.

Population Comprises all those who meet the eligibility criteria (i.e. fulfil the inclusion criteria and exclusion criteria). Not necessarily only the trial participants.

Population summary A well-chosen statistic that allows inference about the population of interest to be derived from a sample of that population.

***p*-value** The probability of observing test results at least as extreme as the result actually observed, under the null hypothesis. This frequentist statistic is typically interpreted as a measure of evidence against the null hypothesis.

Q

Intentionally blank

R

Real-world data Real-world data are data defined by the US Food and Drug Administration (2018) as:

'[...] data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources.' and *Examples of real-world data include data derived from electronic health records; medical claims and billing data; data from product and disease registries; patient-generated data, including in-home-us settings; and data gathered from other sources that can inform on health status, such as mobile devices'*.

Real-world evidence Real-world evidence are defined by the US Food and Drug Administration (2018) as:

'[...] the clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of real-world data'

Regimen Typically used in multi-factorial platform trials to describe a combination of interventions, typically with one intervention per domain.

Registry A collection of information about individuals, for instance patients with a specific diagnosed condition.

Registry trial A pragmatic trial that uses clinical registries as an efficient and low-cost platform for case records, data collection, randomisation, and follow-up.

Randomisation Means by which trial participants are randomly allocated to treatment.

Response adaptive randomisation A randomisation procedure that uses past intervention assignments and participants responses to alter the probability of allocation to different intervention arms/regimens. Typically favours better performing arms/regimens at the time the calculation is made. Can either be 'between-participant' or 'within-participant':

- **Between-participant** Individuals are randomised only once, however that randomisation depends on the observed outcomes of other individuals. Includes group-sequential, multi-arm, multi-stage designs, and response adaptive randomisation.
- **Within-participant** Individuals are randomised, with fixed probability, to multiple treatments sequentially over time, possibly depending on their observed histories. Includes sequential multiple assignment randomised trial designs and some N-of-1 studies.

S

Sample size re-estimation A particular adaptation where the sample size is re-calculated based on data available at an interim analysis. See adaptive design.

Scheduled analysis A pre-specified analysis of trial data, either conditional on elapsed calendar time or some recruitment/follow-up target (e.g. total number of patients with primary outcome), with the intention of either making adaptive trial decisions or providing a terminal analysis. A non-terminal scheduled analysis is also often known as an 'interim analysis'.

Sequential multiple assignment randomised trial (SMART) A trial where each participant is randomised at multiple time-points among a set of interventions that are conditional on the participant's intermediate outcome. See also within-patient randomisation.

Simulation Computer-based experimentation to evaluate operating characteristics of trial designs. Involves simulating large numbers of 'hypothetical' trials with each trial sampling from an assumed probability distribution on, for example, treatment effects, recruitment rates, and times-to-events. Trials designs such as response adaptive designs are typically too complicated to be designed using closed-form mathematical equations. Commonly simulated operating characteristics include type I error (the proportion of false positive successes under the null hypothesis) and power (the proportion of true positive successes under different treatment effects).

Silo An informal shorthand for a group of participants within a platform trial who are defined by some mutually exclusive characteristic.

Stage A time period in an adaptive trial between design adaptations, within which the study design remains fixed.

State A set of mutually exclusive and exhaustive categories, defined by characteristics of a participant within a platform trial, that are capable of changing over time for a single participant at different time-points during their participation in the platform (i.e. they can be dynamic). States are used to define eligibility for domains and this can include defining eligibility that occurs after the time of enrolment. State can used as an additive covariate within the statistical model.

Statistical analysis plan (SAP) A detailed and prespecified instruction for a required analysis of trial data. Often a distinction is made between a terminal SAP (i.e. to be implemented the conclusion of a trial or component of a trial such as a platform domain) or an interim SAP (i.e. to be implemented at a non-terminal scheduled analysis).

Statistical appendix A comprehensive addendum to a core protocol for a platform trial that specifies, in general terms, the randomisation strategies and statistical model/s that will be used to analyse the trial data. Ideally immutable over the course of the trial. This differs from the a SAP in that it will contain only general modelling strategies, whereas a SAP will pro-

vide modelling strategies specific to the data being analysed.

Statistical significance Refers to whether any difference between groups being studied are likely to be real or simply due to chance. It is typically based on pre-specified threshold for the p-value (e.g. 5%).

Strata Comprise a set of mutually exclusive and exhaustive categories (stratum), defined by baseline characteristics of a participant within an adaptive platform trial, in which the relative effects of interventions may be differential. The criteria that define a stratum must be present at or before the time of enrolment. See silos and subgroup.

Subgroup A mutually exclusive baseline population characteristic into which participants of a trial can be partitioned (e.g. adults and non-adults). In the case of a platform trial, a subgroup is some characteristic other than silo membership). See strata and silo.

Superiority A conclusion that an intervention leads to better outcomes than reference treatment.

Stepped wedge A particular type of cluster randomised trial where the intervention is rolled out progressively to all clusters and maintained until the end of the trial. Clusters are randomised in terms of the order in which the intervention is rolled out.

T

Terminal analysis A special case of a scheduled analysis that occurs at the cessation of either a trial, or a component of a trial (e.g. if a domain is closed). The results of such an analysis are typically made publicly available (albeit ensuring the integrity of any remaining components of the trial is not compromised).

Thompson sampling A general framework for sequential decision problems that aims to balance 'exploration' (i.e. learning about the problem) and 'exploitation' (i.e. seeking to maximise expected utility). Also known as posterior sampling. Often used to guide response adaptive randomisation in clinical trials, where randomisation to a treatment is proportional to the probability that the treatment is optimal.

Traditional trial Often used to refer to standard 2 arm, parallel group, randomised clinical trials.

Trial steering committee Also referred to as a 'trial management committee' or 'trial steering group', or similar). Group of investigators, not necessarily independent of the trial, with the overall responsibility for the development and conduct of a trial. This group is typically blinded to treatment allocations until randomisation has ceased.

U

Util A value, where more utils is better. Can be specific, e.g. dollars, or non-specific. See utility and utility function.

Utility Some amount of utils that are accrued upon realising an outcome.

Utility function A function that weights the utilities of a set of outcomes by the probability that the outcome occurs. The inverse of a utility function is known as a loss function. See decision-theoretic, loss function, util, and utility function.

Umbrella trial A special case of a platform trial where there is only a single participant population (i.e. no subgroups), but multiple domains (interventions).

Unit of analysis Within a platform trial the unit of analysis is the group of participants whose data are analysed together within the model for a particular domain. The unit-of-analysis can be all participants who have received an allocation status in that domain or a sub-group of participants who received an allocation status determined by their status with respect to one or more strata or states. Within a domain, the response adaptive randomisation is applied to the unit-of-analysis.

V

Version control Also known as source control. The process of tracking and managing changes to files over time. Essential for trial management. Can refer to either ad hoc methods such as version numbering within documents, or sophisticated software development tools that are increasingly used to transparently manage complicated trial analyses.

Value-of-information Is the amount a decision maker would be willing to pay for information prior to making a decision.

W

Intentionally blank

X

Intentionally blank

Y

Intentionally blank

Z

Intentionally blank

Bibliography

Please note that some of the references included here have not been cited in the main glossary.

Ades AR, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Medical Decision Making 2004;24:207–227.

Angus DC, Berry S, Lewis RJ, Al-Beidh F, Arabi Y, van Bentum-Puijk W, et al. The REMAP-CAP (Randomized Embedded Multifactorial Adaptive Platform for Community-Acquired Pneumonia) Study. Rationale and design. Ann Am Thorac Soc 2020;17:879—891.

Bender R, Stefan L. Adjusting for multiple testing–when and how? Journal of Clinical Epidemiology 2001;54:343–349.

Bell ML, Fiero M, Horton NJ, and Hsu CH. Handling missing data in RCTs; a review of the top medical journals 2014;14:118.

Berger JO. *Statistical Decision Theory*. Springer Series in Statistics. New York: Springer; 1980.

Berry DA. Adaptive clinical trials in oncology. Nature Reviews Clinical Oncology 2012;9:199–207.

Berry SM, Broglio KR, Groshen S, Berry DA. Bayesian hierarchical modeling of patient subpopulations: Efficient designs of Phase II oncology clinical trials. Clinical Trials 2013;10(5):720–734

Berry SM, Carlin BP, Lee JJ, M uller P. Bayesian Adaptive Methods for Clinical Trials. In: Bayesian Adaptive Methods for Clinical Trials 1st ed. Boca Raton, Florida: Chapman and Hall/CRC Press; 2010. p 1–17.

Bofill Roig M, KOnig F, Meyer E, and Posch M. Commentary: two approaches to analyze platform trials incorporating non-concurrent controls with a common assumption. Clinical Trials 2022;19:502–503.

Bofill Roig M, Krotka P, Burman CF, Ekkehard G, Hees K, Jacko P, et al. On model-based time trend adjustments in platform trials with non-concurrent controls. arXiv:211206574v1 [statME] 2021;p 1–37.

Bronson A, CHase MK, Fisher K, Millar D, Perlmutter J, and Nicholas R. Mobilizing the clinical trial ecosystem to drive adoption of master protocols. Clinical Trials 2022;19:690–696.

Bryan J. Excuse me, do you have a moment to talk about version control? The American Statistician 2019;72:20–27.

Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, et al. Stan: a probabilistic programming language. Journal of Statistical Software 2016;20:1–37.

Claxton K, Lacey LF, Walker SG. Selecting treatments: a decision theoretic approach. Journal of the Royal Statistical Society: Series A (Statistics in Society) 2000;163:211–225.

Chakraborty B, and Moodie EEM. Statistical Methods for Dynamic Treatment Regimens. *Statistics for Biology and Health*. New York: Springer; 2013.

Dodd LE, Freidlin B, and Korn EL. Platform trials—beware the noncomparable control group. New England Journal of Medicine 2021;384:1572–1573.

Dunnett CW. A Multiple Comparison Procedure for Comparing Several Treatments with a Control. J Am Stat Assoc 1955;50:1096-1121

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. 3rd ed. Boca Raton, Florida: Chapman and Hall/CRC Press; 2014.

Granholm A, Kaas-Hansen BS, Lange T, Schjørring OL, Andersen LW, Perner A, Jensen AKG, and Moller MH. An overview of methodological considerations regarding adaptive stopping, arm dropping and randomisation in clinical trials. Journal of Clinical Epidemiology 2022:S0895435622002748.

Guyatt GH, Haynes RB, Jaeschke RZ, et al. Users' guides to the medical literature:XXV. evidence-based medicine: principles for applying the users' guides to patient care. JAMA 2000;284(10):1290-1296.

Hee SW, Hamborg T, Day S, Madan J, Miller F, Posch M, Sarah Z, Stallard N. Decision-theoretic designs for small trials and pilot studies: a review. Statistical Methods in Medical Research 2016;25:1022–1038.

Hey SP, Kimmelman J. Are outcome-adaptive allocation trials ethical? Clinical Trials 2015;12:102–106.

Hummel J, Wang S, Kirkpatrick J. Using simulation to optimize adaptive trial designs: applications in learning and confirmatory phase trials. Clinical Investigation 2015;5:401–413.

Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: An assessment of 12 outcomes from 8 studies. Trials 2014;15:139.

Korn EL, Freidlin B. Outcome-adaptive randomization: is it useful? Journal of Clinical Oncology 2011;29:771—776.

Lavori PW, Dawson R. A design for testing clinical strategies: biased adaptive within-subject randomisation. Journal of the Royal Statistical Society: Series A (Statistics in Society) 2000;163:29–28.

Lee KM, Robertson DS, Jaki T, Emsley R. The benefits of covariate adjustment for adaptive multi-arm designs. Statistical Methods in Medical Research 2022;31:2104–2121.

Lee KM, Wason J. Including non-concurrent control patients in the analysis of platform trials: is it worth it? BMC Medical Research Methodology 2020;20:165.

Lipsky AM, and Lewis RJ. Response-adaptive decision-theoretic trial design: operating characteristics and ethics. Statistics in Medicine 2013;32:3752–3765.

Marschner IC, and Schou IM. Analysis of adaptive platform trials using a network approach. Clinical Trials 2022;19:479–489.

Mayer C, Perevozskaya I, Leonov S, Dragalin V, Pritchett Y, Bedding A, Hartford A, Fardipour P, and Cicconetti G. Simulation practices for adaptive trial designs in drus and device development. Statistics in Biopharmaceutical Research 2019;11:325–335.

Meinert, CL. The evolution of data safety and monitoring boards. Clinical Trials 2022:1–3.

McGlothlin AE, Viele K. Bayesian Hierarchical Models. JAMA 2018;320(22):2365–2366.

Meurer W, Lewis R, Berry D. Adaptive clinical trials: a partial remedy for the therapeutic misconception?. Journal of the American Medical Association 2012;307:2377–2378.

Meyer EL, Mesenbrink P, Dunger-Baldauf C, F ülle HJ, Glimm E, Li Y, et al. The Evolution of Master Protocol Clinical Trial Designs: A Systematic Literature Review. Clinical Therapeutics 2020;42:1330–1360

Mirza RD, Punja S, Vohra S, Guyatt G. The history and development of N-of-1 trials. Journal of the Royal Society of Medicine 2017: 110:330–340.

Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods 2019;38:2074–2102.

Murphy SA. Optimal dynamic treatment regimes. Journal of the Royal Statistical Society B 2003;62:331–366.

Murthy S, Fontela P, Berry S. Incorporating Adult Evidence Into Pediatric Research and Practice: Bayesian Designs to Expedite Obtaining Child-Specific Evidence. JAMA 2021;325(19):1937.

Overbey JR, Cheung, YK, and Bagiella E. Integrating non-concurrent controls in the analyses of late-entry experimental arms in multi-arm trials with a shared control group in the presence of parameter drift. Contemporary Clinical Trials 2022; 123:106972

Pallmann P, Bedding AW, Choodari-Oskooei B, Dimairo M, Flight L, Hampson LV, Holmes J, Mander AP, Odondi L, Sydes MR, Villar SS, Wason JMS, Weir CJ, Wheeler GM, Yap C, and Thomas J. Adaptive designs in clinical trials: why use them, and how to run and report them. BMC Medicine 2018;16:1–29.

Park JJH, Detry MA, Murthy S, Guyatt G, Mills EJ. How to Use and Interpret the Results of a Platform Trial: Users' Guide to the Medical Literature. Journal of the American Medicical Assosiation 2022;327:67–74.

Park JJH, Thorland K, and Mills EJ. Critical concepts in adaptive trials 2018;10:343–351.

Park JJH, Siden E, Zoratti MJ, Dron L, Harari O, Singer J, et al. Systematic review of basket trials, umbrella trials, and platform trials: a landscape analysis of master protocols. Trials 2019;20(1):572.

Park JJH, Harari O, Dron L, Lester RT, Thorlund K, Mills EJ. An overview of platform trials with a checklist for clinical readers. Journal of Clinical Epidemiology 2020;125:1–8.

Pike K, Reeves BC, Rogers CA. Approaches to multiplicity in publicly funded pragmatic randomised controlled trials: a survey of clinical trials units and a rapid review of published trials. BMC Medical Research Methodology 2022;22:39.

Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics, 1975; 31:103-115.

Proschan M, and Evans S. Resist the temptation of response-adaptive randomisation. Clinical Infectious Diseases 2020;71:3002–3004.

Robert CP. The Bayesian choice: from decision-theoretic foundations to computational implementation 2nd ed. New York: Springer; 2007.

Robertson DS, Lee KM, López-Kolkovska BC, Villar SS. Response-adaptive randomization in clinical trials: from myths to practical considerations. arXiv:200500564v3 [statME] 2021;p 1–60.

Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. Math Model 1986;7:1393–512.

Ryan EG, Brock K, Gates S, Slade D. Do we need to adjust for interim analyses in a Bayesian adaptive trial design? BMC Medical Research Methodology, 2020;20:150.

Ryan EG, Drovandi CC, McGree JM, Pettitt AN. A review of modern computational algorithms for Bayesian opimal design. International Statistical Review 2016;84:128–154.

Savage LJ. Foundations of Statistics. John Wiley and Sons; 1954.

Saville BR, Berry DA, Berry NS, Viele K, Berry SM. The Bayesian time machine: accounting for temporal drift in multi-arm platform trials. Clinical Trials 2022;19:490–501.

Saville BR, Berry SM. Efficiencies of platform clinical trials: A vision of the future. Clinical Trials. 2016;13(3):358–366.

Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. Trials 2011;12:106.

Senarathne SG, Overstall AM, and McGree JM. Bayesian adaptive N-of-1 trials for estimating population and individual treatment effects. Statistics in Medicine 2020;39:4499–4518.

Simon N, Simon R. Adaptive enrichment designs for clinical trials. Biostatistics 2013;14:613—625.

Stallard N, Todd S, Ryan SG, and Gates S. Comparison of Bayesian and frequentist group-sequential clinical trial designs. BMC Medical Research Methodology 2020;20:4.

Sullivan TR, White IR, Salter AB, Ryan P, Lee KJ. (..) Statistical Methods in Medical Research 2018;27:2610–2626.

Thall PF, Fox P, Wathen J. Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials. Annals of Oncology 2015;26:1621–1628.

Thall PF, and Wathen JK. Practical Bayesian adaptive randomisation in clinical trials. European Journal of Cancer 2007;42:859—866.

Thall PF, and Wathen JK. A simulation study of outcome adaptive randomization. Clinical Trials 14 2017;14:432–440.

Thompson, W. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika 1933;25:285—294.

Thorlund K, Haggstrom J, Park JJH, Mills EJ. Key design considerations for adaptive clinical trials: a primer for clinicians. British Medical Journal 2018;360:k698.

Tsiatis AA, Davidian M, Holloway ST, Laber EB. Dynamic Treatment Regimens. 1st ed. Chapman and Hall/CRC; 2019.

US Food and Drug Administration. Adaptive designs for clinical trials of drugs and biologics. Guidance for industry. Silver Spring, MD: Centre for Drug Evaluation and Research; 2018.

US Food and Drug Administration. Establishment and Operation of Clinical Trial Data Monitoring Committees. Guidance for industry. Silver Spring, MD: Centre for Drug Evaluation and Research; 2006.

U.S. Food Drug Administration. Use of Bayesian Statistics in Medical Device Clinical Trials. Guidance for Industry. Silver Spring, MD: Centre for Drug Evaluation and Research; 2010.

U.S. Food Drug Administration. Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products. Guidance for Industry. Silver Spring, MD: Centre for Drug Evaluation and Research; 2020.

US Food and Drug Administration. Master protocols: efficient clinical trial design strategies to expedite development of oncology drugs and biologics. Guidance for industry. Silver Spring, MD: Centre for Drug Evaluation and Research; 2022.

US Food and Drug Administration. Multiple endpoints in clinical trials. Guidance for industry. Silver Spring, MD: Centre for Drug Evaluation and Research; 2022.

Viele K, Broglio K, McGlothlin A, Saville BR. Comparison of methods for control allocation in multiple arm studies using response adaptive randomization. Clinical Trials 2020;17:52–60

Viele K, Saville BR, McGlothlin A, Broglio K. Comparison of response adaptive randomization features in multiarm clinical trials with control. Pharmaceutical Statistics 2020;19:602–612. .

Villar SS, Robertson DS, Rosenberger WF. The Temptation of Overgeneralizing Response-adaptive Randomization. Clinical Infectious Diseases 2020;p. ciaa1027.

Von Neumann J, Mortgensen O. Theory of games and economic behaviour. Princeton University Press; 1944.

Wang D, Yihan L, Wang X, Liu X, Fu B, Yunzhi L, Larsen L, Offen W. Overview of multiple testing methodology and recent developments in clinical trials. Contemporary Clinical Trials 2015;45:13–20.

Wason JMS, Brocklehurst P, Yap C. When to keep it simple—adaptive designs are not always useful. BMC Medicine 2019;17:152.

Wason JMS, Trippa L. A comparison of Bayesian adaptive randomization and multi-stage designs for multi-arm clinical trials. Statistics in Medicine 2014;33:2206–2221.

Wathen JK, Thall PF. A Simulation Study of Outcome Adaptive Randomization in Multi-arm Clinical Trials. Clinical Trials 2017;14:432–440.

Willan AR, and Pinto EM. The value of information and optimal clinical trial design. Statistics in Medicine 2005;24:1791–1806.

Wilson ECF. A practical guide to value of information analysis. Pharmacoeconomics 2015;33:105–121.

Zucker DR, Schmid CH, McIntosh MW, D'Agostino RB, Selker HP, and Lau J. Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. Journal of Clinical Epidemiology 1997;50:401-410.