Journal of Physiotherapy 65 (2019) 113–116

AUSTRALIAN PHYSIOTHERAPY ASSOCIATION

Journal of PHYSIOTHERAPY

journal homepage: www.elsevier.com/locate/jphys

Appraisal

Research Note: Adaptive trials

The need for innovative clinical trials

The number of clinical trials published each year shows no sign of reaching a plateau,[1] despite the cost of conducting a clinical trial rising five-fold over the past decade.[2] Randomised clinical trials (RCTs) can have a profound and immediate impact on health policy and clinical practice, but while activity and quality have increased, only around half of all trials are published.[3] Even when RCT evidence exists, there may be a failure to adopt this as best practice if the results are considered non-applicable to the specific patient at hand, or where there is mismanagement of commercial and academic interests,[4] or other biases in the design, management, or reporting.[3] Clinicians are often faced with the need to make treatment decisions across a range of comorbidities, whereas evidence from RCTs is predominantly presented for a single medical indication and frequently assumes homogeneity in participant responses. A one-size-fits-all approach to clinical trials is not in harmony with the heterogeneity and complexity of modern diseases. While the pipeline of new treatments is ever-increasing, the capacity to formally evaluate new treatments in RCTs is diminishing, due to the large inherent costs, overburdened healthcare system and low participation rates.[5] Almost half of publicly funded trials do not meet recruitment targets,[6] increasing the costs and the risk of inconclusive results.

Although innovative solutions, such as adaptive designs, exist to improve trial efficiency, trialists rarely invest in the time and resources to simulate a broad range of enrolment and outcome scenarios; this is predominantly due to a lack of statistical expertise in trial simulation and the dearth of simple statistical tools with which to perform this function. Thus, most publicly funded clinical trials have simple parallel group designs (around 90% in the UK) that do not have inbuilt flexibility to adapt to accruing trial evidence.[7] A single conservative stopping boundary applied after 75% of target recruitment could correctly allow early trial termination for futility in up to 30% of trials,[7] saving participants from exposure to futile treatments and decreasing trial costs. Sub-optimal and deleterious treatments continue to be used, and RCTs are increasingly the preserve of large well-resourced pharmaceutical companies. Healthcare-embedded approaches that simultaneously evaluate and implement the best treatment option(s) might provide timely outcomes for participants that are resource-efficient, cost-effective, minimally biased, and accessible to all.

Adaptive trials

The United States of America's Food and Drug Administration (FDA) defines an adaptive trial as 'a clinical trial design that allows for prospectively planned modifications to one or more aspects of the design based on accumulating data from subjects in the trial.'[8] The planned modifications are often defined as thresholds for early stopping due to anticipated trial success or failure. Adaptive methods generally result in changes to the trial parameters, such as the sample size or the ratio of allocation to treatment arms, although more complex adaptations are possible (Box 1). The essential feature is that the rules for all adaptations must be specified in advance in the protocol, and adaptations only occur if a threshold for change has been met due to accruing trial data. The design of these studies typically requires extensive simulations of potential trial outcomes to demonstrate acceptable trial operating characteristics such as the type I (false-positive) error rate.

There are various types of adaptations that can be implemented in a randomised trial. These include: flexible sample size and trial duration, response adaptive randomisation with multiple treatment options and/or doses, adding or dropping of treatments, subgroup evaluations and enrichment, seamless phase 2/3 designs, and adaptive endpoint selection.[9,10] These adaptations can improve trial efficiency and reduce the risk of failed or inconclusive results.[11] Despite the increase in the number of reports of adaptive trials, adopting these innovative designs in publicly-funded clinical trials has been slow. More investment is needed to further develop the statistical methods, software tools, and guidelines for successful implementation.

One reason for the slow adoption of innovative trial designs is that existing competitive mechanisms for publicly funded trials often do not have the flexibility or the timeframe needed to support the necessary planning and simulations required at the design stage, or the finances for ongoing statistical support for repeat interim analyses and trial modifications. Guidelines are needed to aid trialists through the logistics of an adaptive design, and the structure, role, and conduct of Data and Safety Monitoring Boards.[10] Vandemeulebroecke established a framework for the discussion of adaptive designs based on five main points: feasibility, validity, integrity, efficiency, and flexibility.[12]

The relative disadvantages of adaptive trials have been well documented.[10,13,14] Modifications in an ongoing trial, particularly those related to the ratio of allocation to treatment arms, have raised some concerns about the potential to partially unblind or change the type of participants recruited to the study over time. Therefore, additional safeguards are necessary to prevent leakage of information and maintain the integrity of the trial outside the sphere of confidentiality of the Data and Safety Monitoring Board.

Trends in patient populations and in the underlying response to treatment are major issues for all trials. Adaptive designs can be robust to moderate population trends;[13] however, less is known about their sensitivity to time-dependent treatment responses and how accruing information, internal and external to the trial, may be integrated to optimise trial operating characteristics. Many journal articles have provided overviews of adaptive trials, predominantly for drug development and most recently by Bhatt et al,[15] Pallmann et al,[16] and Thorland et al.[17] However, an important and, until now, largely neglected potential for these innovative designs lies within comparative effectiveness trials. Examples of adaptive comparative effective trials include ESETT for status epilepticus,[18] PREPARE ALICE for influenza,[19] and REMAP-CAP for community-acquired pneumonia in patients admitted to intensive care units.[20]

https://doi.org/10.1016/j.jphys.2019.02.004

Box 1. Common adaptive trial designs (may be used in combination).

1. **Sample size re-estimation**
   Adjustments to sample size based on interim parameter estimates such as treatment effect and variance, number lost to follow-up, etc.
2. **Continual assessment (dose-finding)**
   Balances the need to collect sufficient information on each dose, with early stopping rules for treatment arms with poor performance or toxicity.
3. **Response-adaptive randomisation**
   Adjustments to treatment allocation ratios dependent on optimising participant outcome; often proportional to the interim performance of each treatment arm.
4. **Population enrichment**
   Adjustments to trial eligibility criteria to increase recruitment into subgroups that respond well to treatment at interim assessment.

All adaptive designs need to minimise the potential risks associated with: (i) treatment effect estimates being less unreliable in early interim analyses, (ii) leakage of ongoing trial information from multiple interim analyses, (iii) investigator behaviour, participant response or underlying disease changing over time.

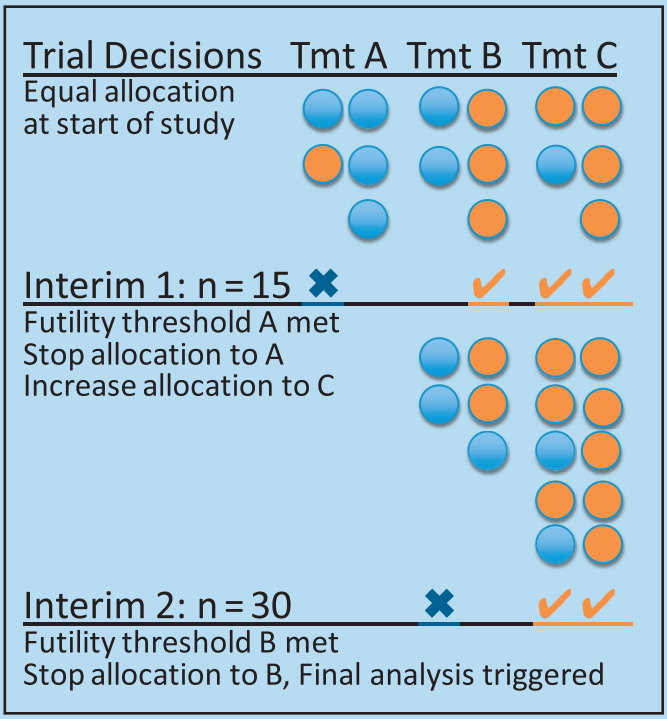
Response adaptive randomisation

Adaptive trials can gain substantial efficiency from response adaptive randomisation (RAR). This is the process where participants are allocated to treatment arms dependent on preceding participant outcomes (Box 2). It has the ethical advantage of assigning fewer participants to poorly performing treatments; however, it can only be implemented when the treatment outcome (or a surrogate) is available within a timely fashion relative to the accrual time required for the entire trial. Concerns have been raised over reductions in statistical power in RAR designs to detect a difference between any two treatment arms (when computed using traditional hypothesis testing procedures) and the deterministic nature of response-adaptive algorithms.[21,22] Solutions have been proposed for multi-arm designs that address these issues by: (i) fixing the allocation to the control group; (ii) using equal allocation to treatment arms until a predefined trial or subgroup size is reached; (iii) defining a minimum probability of allocation to each treatment arm; (iv) applying probabilistic algorithms that apply to blocks of participants; and (v) using post-trial simulations to quantify the bias and provide corrections.[22] Only time trends in the data ('patient drift') of an important magnitude (ie, a change $>$ 25% in the probability of treatment outcome) appear to seriously inflate the risk of a false-positive trial outcome.[23] Response adaptive randomisation designs have been shown to be more efficient than multi-arm multi-stage (MAMS) designs when there is a superior treatment, whereas MAMS designs are slightly more efficient than RAR designs when none of the treatments are effective.[24]

Multi-arm-multi-stage and Bayesian adaptive platform trials

Both MAMS and Bayesian platform trials are concerned with optimising outcomes for a condition rather than focused on any particular therapy.[25,26] The key features of a MAMS trial are the simultaneous comparison of several different treatment options against a single control arm, a weighted randomisation ratio to ensure allocation is discontinued to poorly performing treatment arms, and a frequentist statistical inference and flexible sample size.[25] In 2012, the results for the first MAMS trial – STAMPEDE – an open-label, five-stage, six-arm trial in prostate cancer,[27] were published in *Lancet Oncology*, and identified celecoxib as an inferior option for men with hormone-sensitive and advanced prostate cancer.

Box 2. Representation of response adaptive randomisation in an over-simplified adaptive design.



A total of 30 individuals are recruited over time. After the first 15 individuals have outcome data, the first interim analysis is performed. Treatment A has predominantly poor outcomes (blue dots) and meets the threshold for futility after the first interim analysis (n = 15); therefore, allocation to this arm is stopped. Treatment B has an equal number of good and poor outcomes and it meets the threshold for futility at the second interim analysis (n = 30). Therefore, allocation to Treatment B is also stopped. This triggers the final analysis.

The key features of a Bayesian adaptive platform trial are the use of Bayesian statistical inference and comparison of several different treatment options with innovative adaptive features. Typical adaptations include a flexible sample size, different domains of treatment simultaneously evaluated and in combination (each domain with multiple treatment options), response adaptive randomisation, and evaluation of treatment responses in different subgroups of participants.[26] In 2016, the results for the signature Bayesian platform trial I-SPY 2, a phase 2 trial of neoadjuvant therapy in breast cancer for subgroups of participants defined by genetic signatures, were published and identified better options for women with triple negative and HER2+ breast cancer. [28,29]

Platform trials require extensive consultation between clinicians, consumers, trialists and statisticians in the planning stage, to set the research priorities, and extensive, statistically complex, computer simulations in the design stage. The workload for the trial statistician is considerable during the execution of a trial, which includes multiple interim analyses and potential additional trial simulations to assess planned adaptations. Globally, demand for these statistical skills are a major area of unmet need. These trials also require standardised eligibility criteria, trial endpoints and subgroup definitions, which are documented in a master protocol that is designed to answer multiple research questions.[30] This ensures that treatment responses can be meaningfully aggregated across domains, across sites and over time. The master protocol sets out exactly what data are to be collected, including the primary and secondary endpoints, and the procedures around how data are captured and managed, including the trial governance and safety monitoring arrangements. The trial procedures and analyses for each therapeutic domain are documented in separate appendices. This modular structure allows for the domain-specific appendices to be modified over time without changing the master protocol; treatment options within a domain can be added or removed according to pre-specified adaptation rules, and entire domains can be added or removed. Investment in the development of a master protocol across multiple sites can provide a platform for the ready identification of potentially eligible study participants.

Digital support for adaptive trials

Data management is more challenging for adaptive trials compared to conventional trials, due to the need for timely capture of accurate data for the frequently planned analyses. Participant samples may need to be laboratory analysed and the primary endpoint data made electronically available on an ongoing basis rather than stored and analysed at the end of the trial. Ideally, logic checks need to be automated at the point of data entry and collection tools need to be designed to minimise data entry errors. In common with conventional trials, high standards of data security and privacy must be ensured. The major components of a digital solution are data capture tools, a central or federated database, a treatment allocation tool (that can be updated if RAR is required), an analysis engine (to streamline frequent planned analyses), and a decision engine. Digital solutions that manage the workflow may also be customised to facilitate secure data sharing and provide interfaces to engage with participants, clinicians and study co-ordinators. Lessons learnt from the successful deployment of clinical registries offer insights into the types of digital infrastructure required to implement platform trials, such as online participant registration and consent, and automated notifications and data entry for self-assessments.[31,32] Finally, patient registry platforms can have multi-lingual support.

Bayesian inference and trial simulation

At the heart of an adaptive trial is a decision-making process that occurs in light of the accruing data. Whilst this Research Note is not the place for a full discussion on the use of frequentist or Bayesian inference in adaptive trials, it is worth noting the difference in interpretation of the results between these two approaches. In Bayesian analyses, the level of certainty *in the hypothesis* is represented as a probability, whereas in frequentist analyses, the level of certainty *in hypothetical frequencies of data patterns* is represented as a probability.[33] Bayesian inference does this by combining the *likelihood* of the observation for the range of possible treatment differences with the *prior probability* (ie, prior to data collection) of those possible treatment differences.[34] Bayesian inference provides a straightforward mechanism for updating the estimates of the most probable range of treatment differences as the data accrue.[18] Trial designs that are adaptive can unfold in many different permutations, depending on the accruing data, so it is hard to estimate how 'unlikely' a particular set of results are, making frequentist strategies particularly challenging (though not impossible). Therefore, the ability to update the probabilities for a range of possible treatment differences as new data accrue makes Bayesian inference very useful for adaptive studies.

The ability to generate trial data using simulation is an essential skill when designing adaptive trials. Simulations may be used to design the trial, such as determining the timing and thresholds for repeat analyses, or estimating trial parameters, such as the sample size, or to explore the potential extent of bias in the estimate of treatment effect. For frequentist inference they may also be used to explore whether the confidence intervals have the correct coverage. Some key parameters to consider when performing trial simulations are accrual rate, potential distribution of response in the various treatment arms, number and timing of sequential (interim) analyses, and thresholds for treatment and trial success and futility. In addition, the statistician needs to program an extensive range (often between 50 and 100) of plausible and less plausible trial outcome scenarios, identified by the trial team, to evaluate the trial operating characteristics. Depending on the complexity of the proposed adaptive trial and computer hardware, running a simulation for a single scenario may take anywhere between a few minutes and several days. It is recommended that anyone embarking on a program of adaptive trials starts with simple adaptive designs and builds up experience in the area over time, or engages an experienced statistical team to support the program.

This Research Note has explored the strengths, risks and potential complexity of adaptive trials. The logistical components surrounding the implementation of adaptive trials require a highly integrated, multidisciplinary team encompassing clinicians, statisticians, trialists, consumers, data managers and computer scientists. The challenges are considerable but adaptive trials have been demonstrated to deliver on the promise of simultaneously evaluating and implementing the best treatment option(s) to address the complexity and science of modern diseases.

Competing interests: Nil.

Sources of support: For the manuscript production: JM is supported by a Telethon-Perth Children's Hospital Research Fund research capacity building grant (CRIPTIC); AS holds an NHMRC TRIP Fellowship; BS and SB are supported by a Channel 7 Telethon Trust Major Beneficiary grant; TS holds an NHMRC Career Development Fellowship (no 1111657) and a Raine Clinical Research Fellowship.

Acknowledgements: Nil.

Provenance: Invited. Not peer reviewed.

Correspondence: Dr Julie A Marsh, Telethon Kids Institute, Perth, Australia. Email: Julie.Marsh@telethonkids.org.au

Julie A Marsh[a,b], André Schultz[a,c,d], Benjamin R Saville[e,f] Scott M Berry[e] and Thomas L Snelling[a,g,h,i]

[a]*Wesfarmers Centre of Vaccines & Infectious Diseases, Telethon Kids Institute, Perth, Australia*

[b]*School of Population & Global Health, University of Western Australia, Perth, Australia*

[c]*Faculty of Health and Medical Sciences, University of Western Australia Medical School, Crawley, Perth, Australia*

[d]*Department of Respiratory Medicine, Perth Children's Hospital, Perth, Australia*

[e]*Berry Consultants, Austin, TX, USA*
[f]*Vanderbilt University Department of Biostatistics, Nashville, TN, USA*
[g]*School of Public Health, Curtin University, Bentley, Perth, Australia*
[h]*Department of Infectious Diseases, Perth Children's Hospital, Nedlands, Perth, Australia*
[i]*Menzies School of Health Research and Charles Darwin University, Darwin, Australia*

References

1. Bastian H, et al. *PLoS Med*. 2010;7:e1000326.
2. Collier R. *CMAJ*. 2009;180:277–278.
3. Heneghan C, et al. *BMJ Evid Based Med*. 2017;22:120–122.
4. Bekelman JE, et al. *JAMA*. 2003;289:454–465.
5. McDonald AM, et al. *Trials*. 2006;7:9.
6. Sully BGO, et al. *Trials*. 2013;14:166.
7. Sully BGO, et al. *Trials*. 2014;15:61.
8. *Adaptive Designs for Clinical Trials of Drugs and Biologics Guidance for Industry DRAFT GUIDANCE;* September 2018. https://www.fda.gov/downloads/drugs/guidances/ucm201790.pdf. (accessed 4/2/2019).
9. Berry SM, et al. *Bayesian Adaptive Methods for Clinical Trials*. Florida, USA: CRC Press; 2010.
10. Kairalla JA, et al. *Trials*. 2012;13:145.
11. Saville BR, Berry SM. *Clin Trials*. 2016;13:358–366.
12. Vandemeulebroeke M. *Biometrical J*. 2008;50:541–557.
13. Coffey CS, Kairalla JA. *Drugs R D*. 2008;9:229–242.
14. Berry D. *J Clin Onc*. 2010;29:606–609.
15. Bhatt DL, Mehta C. *N Engl J Med*. 2016;375:65–74.
16. Pallmann P, et al. *BMC Med*. 2018;16:29.
17. Thorlund K, et al. *BMJ*. 2018;360:k698.
18. Connor T, et al. *J Clin Epidemiol*. 2013;66:S130–S137.
19. https://www.phc.ox.ac.uk/phctrials/trial-portfolio/alic4e. (accessed 4/2/2019).
20. https://prepare.ersnet.org/lrmedia/2017/pdf/425.pdf. (accessed 4/2/2019).
21. Hu F, Rosenberger W. *J Am Statist Assoc*. 2003;98:671–678.
22. Villars SS, Rosenberger WF. *Biometrics*. 2018;74:49–57.
23. Villars SS, et al. *Pharm Stat*. 2017;17:1–28.
24. Wason JM, Trippa L. *Stat Med*. 2014;33:2206–2221.
25. Sydes MR, et al. *Trials*. 2009;10:39.
26. Berry SM, et al. *JAMA*. 2015;313:1619–1620.
27. James ND, et al. *Lancet Oncol*. 2012;13:549–558.
28. Rugo HS, et al. *N Engl J Med*. 2016;375:23–34.
29. Park JW, et al. *N Engl J Med*. 2016;375:11–22.
30. Woodcock J, LaVange LM. *N Engl J Med*. 2017;377:62–70.
31. Bellgard M, et al. *Source Code Biol Med*. 2013;8:21.
32. Lacaze P, et al. *Intern Med J*. 2017;47:1075–1079.
33. Greenland S, et al. *Eur J Epidemiol*. 2016;31:337.
34. Dmitrienko A, Wang MD. *Stat Med*. 2006;25:2178–2195.