Please see below for a link to the webinar recording for the Trials Methodology Research Partnership:

**Improving transparency and reporting in adaptive randomised trials**

*Philip Pallman, University of Cardiff and Munya Dimairo, University of Sheffield*

18 September 2020

The slides are also available below.

For any queries, please contact uktnm@nottingham.ac.uk

[https://www.youtube.com/watch?v=Insdfbg5Y3Q](https://www.youtube.com/watch?v=Insdfbg5Y3Q)
Improving transparency and adequate reporting of adaptive randomised trials: introducing the Adaptive designs CONSORT Extension (ACE) guideline

Philip Pallmann
pallmannp@cardiff.ac.uk

Munya Dimairo
m.dimairo@sheffield.ac.uk

MRC-NIHR TMRP Webinar Series
18 September 2020
To adapt or not to adapt?

Being adaptive is a useful thing ...

https://da.wikipedia.org/wiki/Kameleon
Pallmann et al. (2018)
The trials they are a-daptive

Phase II, II/III, and III trials in clinicaltrials.gov (159,645) and the NIHR register (~2300)
The trials they are a-daptive

Adaptive trials in clinicaltrials.gov and ‘the literature’
Adaptive vs. fixed designs

(a) Fixed design analysis plan.

(b) Adaptive design analysis plan.
What is an adaptive design?

A clinical trial design that offers pre-planned opportunities to use accumulating trial data to modify aspects of an ongoing trial while preserving the validity and integrity of that trial. (Dimairo et al. 2018)

→ includes group-sequential and Bayesian methods
→ excludes ‘fully flexible’ designs

Using an adaptive design means...
“...planning to be flexible.” (Shih 2006)
“...taking out insurance.” (Campbell 2013)
“...driving with one’s eyes open.” (Berry 2016)
Embrace flexibility...

... but avoid (excessive) complexity

(a) (b) (c)
Example 1: Group-sequential

**Goal** efficient use of patients, time, and money

**Problem** ignoring clear evidence of futility or efficacy is suboptimal

**Idea** early stopping for futility or efficacy (or safety)

Burnett et al. (2020)
Example 2: Multi-arm multi-stage

**Goal** compare different experimental interventions vs. a reference

**Problem** running multiple controlled trials is inefficient

**Idea** start off with several intervention arms and then drop/select/add
Example 3: Sample size reassessment

**Goal** achieve desired statistical power (e.g. 90%)

**Problem** sample size calculation is often based on vague assumptions

**Idea** get better sample size estimate from interim data
Example 4: Adaptive randomisation

**Goal** compare different interventions

**Problem** subjecting patients to inferior interventions is unethical

**Idea** shift randomisation ratio towards more promising intervention arm

Burnett et al. (2020)
Example 5: Population enrichment

**Goal** focus on patients who benefit most from a treatment

**Problem** not all patients might benefit equally

**Idea** target patients who are most likely to benefit
... and many more

Biomarker adaptive
Adaptive dose ranging
Adaptive hypotheses
Adaptive treatment switching

Seamless phase I/II or II/III
Combinations of several adaptations
Platform, umbrella and basket trials

Syn et al. (2016)
Benefits and challenges

😊 Flexible
- reflects medical practice

😊 Efficient
- shorter trials
- fewer patients
- more accurate estimates

😊 Ethical
- better use of resources
- more patients receive effective treatments

😊 Flexible
- too much flexibility?

😊 Resource intensive
- time to design
- expertise
- software

😊 Complex
- design
- interpretation (bias? post-trial estimation?)
Adaptive designs and bias
Adaptive designs and bias

Statistical bias

Point estimates: over-/underestimation of treatment effects
- use unbiased/bias-adjusted/median-unbiased/shrinkage estimators

Confidence intervals: too wide or too narrow, or ‘mislocated’
- use corrected interval methods

Hypothesis tests: no control of type I error rate (e.g. 5%)
- adjust for multiple (interim) looks at the data
- use bootstrapping and simulations
Adaptive designs and bias

Operational bias

Knowledge (or mere speculative knowledge) of interim results may alter the behaviour of trial investigators and participants.

Inconsistencies in the conduct of the trial before and after interim analyses could lead to heterogeneity in patient population.
Minimising operational bias

Some key questions:

- Who knew what or had access to what, and when?
- Who was to make interim decisions?
- What was the sponsor’s role in interim decision making?
- (How) were confidentiality and blinding maintained?
- Is the patient population likely to be heterogeneous across trial stages?

https://kimbia.com/crowdfunding-right-3-key-questions-organizations-answer/
Minimising operational bias

In an ideal world . . .

✓ all possible adaptations are laid down in advance
✓ blinding is maintained wherever possible
✓ only an independent statistician and DMC are allowed to see interim data
✓ all adaptations made are described
✓ both unbiased/bias-adjusted and standard (maximum likelihood) estimates are reported
✓ (worst-case) bias is quantified e.g., via simulation
✓ tests and confidence intervals are adjusted for multiple looks
Now we are familiar with:

• What adaptive designs are
• Examples of trial adaptations
• Statistical and non-statistical issues that can introduce bias

There are additional transparency and reporting demands
We don’t just need to report adaptive trials but we need to report them well

1. Difficult to reproduce methods and results
2. Difficult to interpret their results
3. Hard to synthesize evidence
4. Hamper the ability of their results to inform practice
5. Limit their contribution to future research
6. Contribute to research waste

Inconsistent and inadequate reporting of adaptive design clinical trials
How we developed the reporting guideline

Development process of a consensus-driven CONSORT extension for randomised trials using an adaptive design

Munyaradzi Dimairo1, Elizabeth Coates1, Philip Pallmann2, Susan Todt3, Steven A. Julious1, Thomas Jaki4, James Wason5,6, Adrian P. Mander7, Christopher J. Wei8, Franz Koenig7, Marc K. Walton8, Katie Biggs3, Jon Nicholl9, Toshimitsu Hamasaki9, Michael A. Proschan10, John A. Scott11, Yuki Ando12, Daniel Hind1 and Douglas G. Altman1

Abstract

Background: Adequate reporting of adaptive designs (ADs) maximises their potential benefits in the conduct of clinical trials. Transparent reporting can help address some obstacles and concerns related to the use of ADs. Currently, there are deficiencies in the reporting of AD trials. To overcome this, we have developed a consensus-driven extension to the CONSORT statement for randomised trials using an AD. This paper describes the processes and methods used to develop this extension rather than detailed explanation of the guideline.
ACE guideline

The BMJ (https://www.bmj.com/content/369/bmj.m115)
CONSORT website (http://www.consort-statement.org/extensions/overview/randomised-trials-using-adaptive-designs)
EQUATOR Network website (https://www.equator-network.org/reporting-guidelines/the-adaptive-designs-consort-extension-ace-statement/)
Other platforms
What is covered in the ACE E&E

• Some background (what is an adaptive design, examples of trial adaptations, current use and reporting issues...)

• Checklist items (abstract and main report), why they are important, what is expected when reporting, examples

• Checklists to download and complete when submitting your manuscript
Scope/principles of the ACE guideline

• Randomised trials using adaptive designs
• Not for internal pilots assessing operational feasibility
• Statistical paradigm used doesn’t matter (frequentist or Bayesian)
• Future proof (generic)
• Minimum essential requirements
• It’s about access to information (not where information should be)
• Used alongside other extensions when appropriate
Contents

• Main checklist
  – 7 new items,
  – 9 modified items,
  – 6 unchanged items with additional explanatory text

• Abstract checklist
  – 1 new item
  – 1 modified item
  – 1 unchanged item with additional explanatory text
We strongly encourage researchers to use the checklists cross-referencing the detailed E&E statement.

It’s not a tick box exercise!!

Use it throughout the trial from the design stage!
ACE Abstract (modification) – “Trial design”

- Description of the trial design (for example, parallel, cluster, non-inferiority); include the word ‘adaptive’ in the content or at least as a keyword.

Rationale
- Better indexing of adaptive trials so others can retrieve them easily
- Study design may influence interpretation of results and evidence synthesis approach

Specific expectations
- It does not matter which part of the abstract
- May wish to state pre-planned trial adaptations (if possible)

Example: “AWARD-5 was an adaptive, seamless, double-blind study comparing dulaglutide, a once-weekly glucagon-like peptide-1 (GLP-1) receptor agonist, with placebo at 26 weeks and sitagliptin up to 104 weeks.” and keyword “Bayesian adaptive”’
ACE item 3b (new): Pre-planned adaptive design features

- Type of adaptive design used, with details of the pre-planned trial adaptations and the statistical information informing the adaptation

**Rationale**

- Indicate design concepts and applicable statistical methods
- Pre-specification is essential for trial credibility
- Important for assessing appropriateness of statistical methods used
- Reproducibility and interpretation of results

**Specific expectations**

- Leave no room for ambiguity
- Scientific rationale for considering the trial adaptations (linked to CONSORT 2010 item 2a)
- Statistical models or formulae for gathering statistical information to guide trial adaptations

**Example 3. Pre-planned adaptations; 5-arm 2-stage AD allowing for regimen selection, early stopping for futility and SSR**

- “This randomized, placebo-controlled, double-blind, phase 2/3 trial had a two-stage adaptive design, with selection of the propranolol regimen (dose and duration) at the end of stage 1 (interim analysis) and further evaluation of the selected regimen in stage 2. Pre-specified possible adaptations to be made after the interim analysis, as outlined in the protocol and statistical analysis plan (accessible via journal website), were selection of one or two regimens, sample-size reassessment, and non-binding stopping for futility.”
ACE item 6a (*modification*)

- Completely define pre-specified primary and secondary outcome measures, including how and when they were assessed. *Any other outcome measures used to inform pre-planned adaptations should be described with the rationale.*

**Rationale**

- Some adaptations may be based on quickly observed outcomes (adaptation outcomes) other than the primary and secondary outcomes
- Adaptation outcomes influence adaptation process, statistical characteristics, clinical interpretation and trustworthiness of results
- Ability to judge the reliability of adaptation outcomes and related adaptations made

**Specific expectations**

- Rationale to support that adaptation outcomes are reliable

*Example 4. MAMS AD; adaptation rationale (part of item 3b); rationale for adaption outcome different from the primary outcome; description of the adaptation and primary outcomes*

  "This seamless phase 2/3 design starts with several trial arms and uses an intermediate outcome to adaptively focus accrual away from the less encouraging research arms, continuing accrual only with the more active interventions. The definitive primary outcome of the STAMPEDE trial is overall survival (defined as time from randomisation to death from any cause). The intermediate primary outcome is failure-free survival (FFS) defined as the first of: PSA failure (PSA >4 ng/mL and PSA >50% above nadir); local progression; nodal progression; progression of existing metastases or development of new metastases; or death from prostate cancer. FFS is used as a screening method for activity on the assumption that any treatment that shows an advantage in overall survival will probably show an advantage in FFS beforehand, and that a survival advantage is unlikely if an advantage in FFS is not seen. Therefore, FFS can be used to triage treatments that are unlikely to be of sufficient benefit. It is not assumed that FFS is a surrogate for overall survival; an advantage in FFS might not necessarily translate into a survival advantage."167
ACE item 7b (replacement)

- Pre-planned interim decision-making criteria to guide the trial adaptation process; whether decision-making criteria were binding or non-binding; pre-planned and actual timing and frequency of interim data looks to inform trial adaptations

Rationale

- Influence operating characteristics, reliability of the adaptations made, interpretation and credibility of results
- Ability to judge the implications of overruling or ignoring adaptation decision rules

Specific expectations

- Decision rules describing how and when the proposed adaptations will be made
- Criteria for claiming overall evidence
ACE item 8b (*modification*)

- Type of randomisation; details of any restriction (such as blocking and block size); any changes to the allocation rule after trial adaptation decisions; any pre-planned allocation rule or algorithm to update randomisation with timing and frequency of updates

**Rationale**

- Allocation ratios can be fixed throughout, updated as an adaptation, or changed as a result of adaptations or unplanned changes made

- Important for *response adaptive randomization* as it influence design efficiency, operating characteristics, and trustworthiness of results

- No unique approach to update randomization

**Specific expectations**

- Burn-in period before adaptive randomisation

- Algorithms used for adaptive randomisation

- Decision-making criteria for stopping arms or trial (if applicable)
ACE item 11c (new): Confidentiality and minimisation of operational bias

- Measures to safeguard the confidentiality of interim information and minimise potential operational bias during the trial

**Rationale**

- Additional sources of bias due to access to interim data/results (or mere speculation)
- May cause inconsistencies in trial conduct (e.g., clinical management before and after adaptation)
- Combining data across stages may be questionable
- Complicates interpretation of results
- Not disclosed in most adaptive trials

**Specific expectations**

- Who had access to interim data and performed analyses
- How confidentiality was safeguarded (e.g., communication process)
- Adaptation decision-making process and roles of key stakeholders (e.g., sponsor)

See detailed examples in the guidance document
ACE Item 12b (new): Estimation and inference methods

- For the implemented adaptive design features, statistical methods used to estimate treatment effects for key endpoints and to make inferences

**Specific expectations**

- Methods during interim and final analyses
- Methods for combining data across stages and weights used and for controlling stated operating characteristics
- Whether simulation was used (e.g., to evaluate magnitude of bias) and simulation report (item 24c)
- For Bayesian methods, models for estimating posterior probabilities; priors used and rationale for its choice; whether priors were updated using interim data and how; sources of data for informative priors (when applicable)

---

**Example 3. 3-arm 2-stage group-sequential AD with treatment selection; combination test method; multiplicity adjustments; statistical method for estimating treatment effects**

- “The proposed closed testing procedure will combine weighted inverse normal combination tests using pre-defined fixed weights, the closed testing principle,\(^{68, 253, 254}\) and the Hochberg-adjusted 1-sided P-value on stage 1 data. This testing procedure strongly controls the overall type I error rate at a level (see “Simulations run to assess the type I error rate under several null hypothesis scenarios”). Multiplicity-adjusted flexible repeated 95% 2-sided CIs\(^{217}\) on the percentage of patients will be calculated for otamixaban dose 1, otamixaban dose 2, and UFH plus eptifibatide. Relative risk and its 95% 2-sided CIs will also be calculated. Point estimates based on the multiplicity-adjusted flexible repeated CIs will be used.\(^{181}\) See supplementary material of the paper for details.
ACE item 13a (modification) – Participant flowchart

- For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome and any other outcomes used to inform pre-planned adaptations, if applicable.

**Specific expectations**

- Applies to both interim and final analyses depending on the stage of reporting.
- Flowchart should be consistent with key hypotheses (e.g., reflect subpopulations and full populations if applicable).
- Participants with adaptation outcome data (per group) if different from the primary outcome.
- Participants who did not contribute to interim analysis, with reasons (e.g., immature outcome data).
ACE Item 14c (new): Adaptation decisions

Specify what trial adaptation decisions were made in light of the pre-planned decision-making criteria and observed accrued data

Rationale

- Essential to adhere to pre-planned decision rules to inform trial adaptations
- Deviations from pre-planned decision rules may occur (unforeseeable events)
- Interim decisions vs pre-planned decisions are poorly reported

Specific expectations

- When decision were made (timing)
- What adaptations were enforced; not enforced or overruled
- Unplanned changes that were made with reasons
- Design changes as a consequence of made adaptation decisions

Example 2. Dose-selection decisions for an inferentially seamless phase 2/3 AD

“The two doses of indacaterol selected against the two reference efficacy criteria were 150 µg (as the lowest dose exceeding both criteria) and 300 µg (as the next highest dose). The safety results, together with the safety data from the other 1-year study, led the DMC to conclude that there was no safety signal associated with indacaterol at any dose. Thus, the two doses selected (at stage 1) to continue into stage 2 of the study were indacaterol 150 and 300 µg.”
ACE Item 15b (new): Similarity between stages

Summary of data to enable the assessment of similarity in the trial population between interim stages

Rationale

- Patient characteristics and standard of management may change before and after trial adaptations
- Results may be inconsistent between stages complicating interpretation of results vs intended objectives
- Ability to assess similarity in trial population between stages and consistency between groups

Specific expectations

- Overall characteristics by stage
- Characteristics by group at each stage

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Characteristics of randomised participants (N=1202) in stage 1 and 2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Stage 1 (n=230)</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>53.4 (10.3)</td>
</tr>
<tr>
<td>Gender (female), n (%)</td>
<td>139 (60.4)</td>
</tr>
<tr>
<td>Race (white), n (%)</td>
<td>103 (44.8)</td>
</tr>
<tr>
<td>BMI (kg/m²), mean (SD)</td>
<td>31.9 (4.5)</td>
</tr>
<tr>
<td>Body weight (kg), mean (SD)</td>
<td>87.3 (18.0)</td>
</tr>
<tr>
<td>Duration of diabetes (years), mean (SD)</td>
<td>7.5 (5.5)</td>
</tr>
<tr>
<td>Seated systolic BP (mm Hg), mean (SD)</td>
<td>128.0 (14.4)</td>
</tr>
<tr>
<td>Seated diastolic BP (mm Hg), mean (SD)</td>
<td>77.9 (7.9)</td>
</tr>
<tr>
<td>Seated heart rate (bpm), mean (SD)</td>
<td>74.5 (9.6)</td>
</tr>
</tbody>
</table>

Adapted from Geiger et al.66; BMI = Body Mass Index; SD = standard deviation; BP = blood pressure; bpm = beats per minute; mm Hg = millimetres of mercury. Data presented were from an ongoing trial so are incomplete and only used for illustration.
ACE Item 17c (new): Interim results

- Report interim results used to inform interim decision-making

**Rationale**

- To judge whether pre-planned trial adaptations and decision rules were adhered to
- To assess consistency in results across stages

**Specific expectations**

- Results/data used to inform each pre-planned trial adaptation (3b) – linked to decision-making criteria (7b)
- All treatment groups or subpopulations including those stopped early (e.g., for futility)

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Interim results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parasite clearance at day 30 (initial cure)</strong></td>
<td><strong>Treatment group</strong></td>
</tr>
<tr>
<td>Interim analysis 1</td>
<td>Single dose, 7.5mg/kg</td>
</tr>
<tr>
<td></td>
<td>Multiple dose, 7x3mg/kg</td>
</tr>
<tr>
<td>Interim analysis 2</td>
<td>Single dose, 10mg/kg</td>
</tr>
<tr>
<td></td>
<td>Multiple dose, 7x3mg/kg</td>
</tr>
<tr>
<td>Interim analysis 3</td>
<td>Single dose, 10mg/kg</td>
</tr>
<tr>
<td></td>
<td>Multiple dose, 7x3mg/kg</td>
</tr>
</tbody>
</table>

N, total number of patients per group (denominator); n, patients with recorded parasitic clearance per groups (events); CI, confidence interval; *p-value from Fisher’s exact test, adaptation rule met to escalate dose so dosage increased to 10 mg/kg and continue recruitment; adaptation rule to escalate dose not met so recruitment was continued with the same dosage (10mg/kg in single-dose arm; values from a Chi-square test; adaptation rule to escalate dose not met but concerns arose regarding low cure in each arm and recruitment was terminated; includes patients in interim analysis 2; patients in interim analysis 1 did not contribute to any subsequent interim analysis.
ACE Item 24b (new): Statistical analysis plan & other related trial documents

Where the full statistical analysis plan and other relevant trial documents can be accessed

**Rationale**
- Contains detailed statistical methods
- Critical details of the trial adaptations may be intentionally withheld when the trial is ongoing to minimize operational bias
- Details of statistical simulation and report may be required
- Transparency regarding decision-making process, roles and responsibilities of those involved, and recommendations made

**Specific expectations**
- Latest SAP versions for both interim and final analyses, including amendments
- Simulation details, if applicable (e.g., simulation protocol/plan and report of results or related publications)
- Data monitoring/trial adaptation committee charter and recommendations made
Conclusions

• We all have a part to play to improve transparency and adequate reporting of adaptive trials
• ACE is just one piece of the puzzle
• More needs to be done to train multidisciplinary stakeholders (e.g. researchers, journal editors, reviewers)
Acknowledgments

• Funders: NIHR CTU Support Funding scheme and MRC HTMR
• ACE Steering Committee
• ACE Consensus Group
• ACE External Expert Panel
• Many who contributed during the development process
References


