Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Adaptive designs CONSORT Extension (ACE) Statement: A checklist with explanation and elaboration guideline for reporting randomised trials that use an adaptive design

Munyaradzi Dimairo (m.dimairo@sheffield.ac.uk)
University of Sheffield https://orcid.org/0000-0002-9311-6920
Philip Pallmann
Centre for Trials Research, Cardiff University
James Wason
Institute of Health and Society, Newcastle University
Susan Todd
University of Reading
Thomas Jaki
Lancaster University
Steven A. Julious
School of Health and Related Research, University of Sheffield
Adrian P. Mander
MRC Biostatistics Unit, University of Cambridge
Christopher J. Weir
Usher Institute of Population Health Sciences and Informatics, University of Edinburgh
Franz Koenig
Centre for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna
Marc K. Walton
Janssen Pharmaceuticals
Jon P. Nicholl
School of Health and Related Research, University of Sheffield
Elizabeth Coates
School of Health and Related Research, University of Sheffield
Katie Biggs
School of Health and Related Research, University of Sheffield
Toshimitsu Hamasaki
National Cerebral and Cardiovascular Center
Michael A. Proschan
National Institute of Allergy and Infectious Diseases, National Institutes of Health
John A. Scott
Division of Biostatistics in the Centre for Biologics Evaluation and Research, Food and Drug Administration
Yuki Ando
Pharmaceuticals and Medical Devices Agency
Daniel Hind
School of Health and Related Research, University of Sheffield
Douglas G. Altman
Centre for Statistics in Medicine, University of Oxford

Methodology

Keywords: Adaptive design; flexible design; transparency; CONSORT extension; reporting guideline; Randomised trial

Posted Date: May 21st, 2019

DOI: https://doi.org/10.21203/rs.2.9725/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published on June 17th, 2020. See the published version at https://doi.org/10.1186/s13063-020-04334-x.

Abstract

Background Adaptive designs (ADs) allow pre-planned changes to an ongoing trial without compromising the validity of conclusions and it is essential to distinguish pre-planned from unplanned changes that may also occur. The reporting of ADs in randomised trials is inconsistent and needs improving. Incompletely reported AD randomised trials are difficult to reproduce and are hard to interpret and synthesise. This consequently hampers their ability to inform practice as well as future research and contributes to research waste. Better transparency and adequate reporting will enable the potential benefits of ADs to be realised. Methods This extension to the Consolidated Standards Of Reporting Trials (CONSORT) 2010 Statement was developed to enhance the reporting of randomised AD clinical trials. We developed an Adaptive designs CONSORT Extension (ACE) guideline through a two-stage Delphi process with input from multidisciplinary key stakeholders in clinical trials research in the public and private sectors from 21 countries, followed by a consensus meeting. Members of the CONSORT Group were involved during the development process. Results The paper presents the ACE checklists for AD randomised trial reports and abstracts, as well as an explanation with examples to aid the application of the guideline. The ACE checklist is comprised of seven new items, nine modified items, six unchanged items for which additional explanatory text clarifies further considerations for ADs, and 20 unchanged items not requiring further explanatory text. The ACE abstract checklist has one new item, one modified item, one unchanged item with additional explanatory text for ADs, and 15 unchanged items not requiring further explanatory text. Conclusions The intention is to enhance transparency and improve reporting of AD randomised trials to improve the interpretability of their results and reproducibility of their methods, results and inference. We also hope indirectly to facilitate the much-needed knowledge transfer of innovative trial designs to maximise their potential benefits.

Purpose Of The Paper

Incomplete and poor reporting of randomised clinical trials makes trial findings difficult to interpret due to study methods, results, and inference that are not reproducible. This severely undermines the value of scientific research, obstructs robust evidence synthesis to inform practice and future research, and contributes to research waste [1,2]. The Consolidated Standards Of Reporting Trials (CONSORT) statement is a consensus-based reporting guidance framework that aims to promote and enhance transparent and adequate reporting of randomised trials [3,4]. Specific CONSORT extensions addressing the reporting needs for particular trial designs, hypotheses, and interventions have been developed [5]. The use of reporting guidelines is associated with improved completeness in study reporting [6–8]; however, mechanisms to improve adherence to reporting guidelines are needed [9–12].

We developed an Adaptive designs CONSORT Extension (ACE) [13] to the CONSORT 2010 statement [3,4] to support reporting of randomised trials that use an adaptive design (AD) – referred to as AD randomised trials. In this paper, we define an AD and summarise some types of ADs as well as their use and reporting. We then describe briefly how the ACE guideline was developed; present its scope and underlying principles. Finally, we present the ACE checklist with explanation and elaboration (E&E) to guide its use.

Adaptive designs: definition, current use and reporting

The ACE Steering Committee [13] agreed a definition of an AD (*Box 1*) consistent with the literature [14–18].

Box 1. Definition of an AD

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.

Substantial uncertainties often exists when designing trials around aspects such as the target population, outcome variability, optimal treatments for testing, treatment duration, treatment intensity, outcomes to measure, and measures of treatment effect [19]. Well designed and conducted AD trials allow researchers to address research questions more efficiently by allowing key aspects or assumptions of ongoing trials to be evaluated or validly stopping treatment arms or entire trials on the basis of available evidence [15,18,20,21]. As a result, patients may receive safe, effective treatments sooner than with fixed (non-adaptive) designs [19,22–25]. Despite their potential benefits, there are practical challenges and obstacles to the use of ADs [18,26–33].

The literature on ADs is considerable and there is specific terminology associated with the field. *Box 2* gives a glossary of key terminology used throughout this E&E document.

Box 2. Definitions of key technical terms

Validity	The ability to provide correct statistical inference to establish effects of study interventions and produce accurate estimates of effects (point estimates and uncertainty), to give results that are convincing to the broader audience (science community and consumers of research findings).
Integrity	Relates to minimisation of operational bias, maintenance of data confidentiality, and ensuring consistency in trial conduct (before and after adaptations) for credibility, interpretability, and persuasiveness of trial results.
Pre-planned adaptations or adaptive features	Pre-planned/prespecified changes or modifications to be made to aspects of an ongoing trial, which are specified at the design stage or at least before seeing accumulating trial data by treatment group, and are documented for audit trail (e.g. in the protocol).
Unplanned changes	Ad hoc modifications to the aspects of an ongoing trial.
Type of AD	The main category used to classify a trial design by its pre-planned adaptive features or adaptations. Some ADs can fall into more than one main category of trial adaptations.
Adaptive decision-making criteria	Elements of decision-making rules describing *whether, how* and *when* the proposed trial adaptations will be utilised during the trial. It involves pre-specifying a set of actions guiding how decisions about implementing the trial adaptations are made given interim observed data (decision rules). It also involves pre-specifying limits or parameters to trigger trial adaptations (decision boundaries). For example, stopping boundaries that relate to pre-specified limits regarding decisions to stop the trial or treatment arm(s) early.
Interim analysis	Statistical analysis or review of accumulating data from an ongoing trial (interim data) to inform trial adaptations (before the final analysis), which may or may not involve treatment group comparisons.
Binding rules	Decision rules that must be adhered to for the design to control the false-positive error rate
Non-binding rules	Optional decision rules that can be overruled without negative effects on control of the false-positive error rate
Statistical properties or operating characteristics	Relates to behaviour of the trial design. These may include statistical power, false-positive error rate, bias in estimation of treatment effect(s), or probability of each adaptation taking place
Simulation	A computational procedure performed using a computer program to evaluate statistical properties of the design by generating pseudo data according to the design, under a number of scenarios and repeated a large number of times
Fixed (non-adaptive) design	A clinical trial that is designed with an expected fixed sample size without any scope for pre-planned changes (adaptations) of any study design feature
Bias	The systematic tendency for the treatment effect estimates to deviate from their 'true values' including the statistical properties (e.g., error rates) to deviate from what is expected in theory (e.g., pre-specified nominal error rate)
Operational bias	Occurs when knowledge of key trial-related information influences changes to the conduct of that trial in a manner that biases the conclusions made regarding the benefits and/or harms of study treatments
Statistical bias	Bias introduced to the study results or conclusions by the design such as a result of changes to aspects of the trial or multiple analysis of accumulating data from an ongoing trial
Subpopulation(s)	Subset(s) of the trial population that can be classified by characteristics of participants that are thought to be associated with treatment response (e.g., genetic markers or biomarkers)
Adaptation outcome(s)	Outcome(s) used to guide trial adaptation(s); they may be different from the primary outcome(s)

Box 3 summarises some types of ADs and cites examples of their use in randomised trials. The motivations for these trial adaptations are well discussed [15,18,22,25,34–37]. Notably, classification of ADs in the literature is inconsistent [13,22], while the scope and complexity of trial adaptations and underpinning statistical methods continues to broaden [18,20,38].

Box 3. Some types of adaptations used in randomised trials with examples

Trial adaptive feature or adaptation, motivation, and cited examples of use	Type of AD and examples of underlying statistical methods
Changing the predetermined sample size in response to inaccurate assumptions of study design parameters to achieve the desired statistical power [39–41].	Sample size re-estimation/re-assessment/re-calculation (SSR) using aggregated interim data from all participants or interim data separated according to allocated treatment [42–49].
Stopping the trial early for efficacy, futility or safety when there is sufficient evidence [50,51].	Group sequential design (GSD) [52,53]; information-based GSD [54]; futility assessment using stochastic curtailment [55–57].
Evaluating multiple treatments in one trial allowing for early selection of promising treatments or dropping futile or unsafe treatments [58–60]. New treatments can also be added to an ongoing trial [61].	Multi-arm multi-stage (MAMS), dose/treatment-selection, drop-the-loser, or pick-the-winner, or add arm [23,62–72].
Changing the treatment allocation ratio to favour treatments indicating beneficial effects [73,74].	Response-adaptive randomisation (RAR) [75–80].
Investigating multiple research objectives that are traditionally examined in distinct trial phases, in one trial under a single protocol [81–83]. For instance, addressing learning (selecting promising treatments for further testing) and confirmatory objectives in one trial.	Operationally or inferentially seamless AD [69,71,84–87].
Adjusting the trial population or selecting patients with certain characteristics that are most likely to benefit from investigative treatments [88–91]. This may involve incorporating statistical information from or adapting on a biomarker.	Population/patient enrichment or biomarker AD [92–96].
Changing the primary research hypotheses/objectives or primary endpoints [86,97]. For example, switching from non-inferiority to superiority.	Adaptive hypotheses [98,99].
Switching the allocated treatment of patients to an alternative treatment influenced by ethical considerations, for instance, due to lack of benefit or safety issues.	Adaptive treatment-switching [100,101].
Combination of at least two types of adaptations [24,41,97,102–107].	Multiple ADs such as GSD or drop-the-loser with SSR [108]; inferentially seamless phase 2/3 AD with hypotheses selection [85] or population enrichment [109]; biomarker-stratified with RAR [110]; adaptive platform trials where arms can be added or stopped early [19,24,111].

Furthermore, there is growing literature citing AD methods [29,86,112] and interest in their application by researchers and research funders [26,28,113]. Regulators have published reflection and guidance papers on ADs [14,113–116]. Several studies, including regulatory reviews, have investigated the use of ADs in randomised trials [27,29,31,33,42,106,112,113,117–125]. In summary, ADs are used in a relatively low proportion of trials, although their use is steadily increasing in both the public and private sectors [119–121], and they are frequently considered at the design stage [27].

The use of ADs is likely to be underestimated due to poor reporting making it difficult to retrieve them in the literature [119]. Whilst the reporting of standard CONSORT requirements of AD randomised trials is generally comparable to that of traditional fixed designed trials [117], inadequate and inconsistent reporting of essential aspects relating to ADs is widely documented [26,27,50,112,118,125–128]. This may limit their credibility, the interpretability of results, and their ability to inform or change practice [14,26–28,30,31,113,114,124–126], whereas transparency and adequate reporting can help address these concerns [22,27]. In summary, statistical and non-statistical issues arise in ADs [22,37,106,113,129–134], which require special reporting considerations [13].

Summary of how the ACE guideline was developed

We adhered to a consensus-driven methodological framework for developing healthcare reporting guidelines recommended by the CONSORT Group and EQUATOR Network [135]. A registered protocol was developed [136] and a paper detailing the rationale and the development process are publicly accessible [13]. This comprised of a two-stage Delphi process involving international multidisciplinary key stakeholders in clinical trials research employed in the private and public sectors, which was followed by a consensus meeting. Members of the CONSORT Group provided oversight throughout. The ACE Consensus Group and Steering Committee approved the final checklist and contributed to this E&E document. *Box 4* outlines the scope of principles guiding the application of this extension.

Box 4. ACE guideline scope and general principles

1. It applies to all randomised clinical trials using an AD, as defined in *Box 1*.

2. It excludes randomised clinical trials that change aspects of an ongoing trial based entirely on external information [137] or with internal pilots focusing solely on feasibility and processes (e.g. recruitment, intervention delivery and data completeness) [138].

3. It covers general reporting principles to make it applicable to a wide range of current and future ADs and trial adaptations.

4. It is not intended to promote or discourage the use of any specific type of AD, trial adaptation, or frequentist or Bayesian statistical methods. These choices should be driven by the scientific research questions, the goals behind the use of the proposed AD features, and practical considerations [22].

5. It aims to promote transparent and adequate reporting of AD randomised trials to maximise their potential benefits and improve the interpretability of their results and their reproducibility, without impeding their appropriate use or stifling design innovation. Therefore, the guideline does not specifically address the appropriateness of adaptive statistical methods.

6. It presents the minimum requirements that should be reported but we also encourage authors to report additional information that may enhance the interpretation of trial findings.

7. Access to information is most important regardless of the source and form of publication. For example, use of appendices and citation of accessible material (e.g. protocols, statistical analysis plans (SAPs), or related publications) is often sufficient.

8. The order in which researchers report information does not necessarily need to follow the order of the checklist.

9. The guideline does not primarily address specific reporting needs for non-randomised ADs (e.g. phase I dose escalation studies, phase II single-arm designs). However, some principles covered here may still apply to such trials.

Structure of the ACE guideline

Authors should apply this guideline together with the CONSORT 2010 statement [3,4] and any other relevant extensions depending on other design features of their AD randomised trial. *Box 5* summarises the changes made to develop this extension. We provide both CONSORT 2010 and ACE items with comments, explanation, and examples to illustrate how specific aspects of different types of AD randomised trials should be reported. For the examples, we obtained some additional information from researchers or other trial documents (e.g., SAPs and protocols). Headings of examples indicate the type of AD and the specific elements of an item that were better reported, so examples may include some incomplete reporting in relation to other elements.

Box 5. Summary of significant changes to the CONSORT 2010 statement

New items: Introduces seven new items that are specific to AD randomised trials	§ 3b on pre-planned AD features,§ 11c on confidentiality and minimisation of operational bias,§ 12b on estimation and inference methods,§ 14c on adaptation decisions,§ 15b on similarity between stages,§ 17c on interim results and,§ 24b on SAP and other relevant trial documents.
Restructuring: Renumbers four standard items to accommodate the new items	§ 3b is now 3c (on losses and exclusions) to accommodate the new item 3b,§ 12b is now 12c (on methods for additional analyses) to accommodate the new item 12b,§ 15 on baseline demographics and clinical characteristics is now 15a to accommodate new item 15b and,§ 24 on access to protocol is now 24a to accommodate new item 24b.
Modified items: Modifies nine standard items	§ 3b (now 3c) on important changes to the design or methods after commencement,§ 6a on pre-specified primary and secondary outcomes,§ 6b on changes to trial outcomes after commencement,§ 7a on sample size,§ 7b on interim analyses and stopping rules, which is now a replacement capturing adaptive decision-making criteria to guide adaptation(s),§ 8b on type of randomisation,§ 12a on statistical methods to compare groups,§ 13a on participants randomised, treated and analysed§ 14a on dates for recruitment and follow-up.
Expanded text: Expands the E&E text for clarification on six items without changes to item wording	§ 14b on why the trial ended or was stopped,§ 15 (now 15a) on baseline demographics and clinical characteristics),§ 16 on numbers randomised,§ 17a on primary and secondary outcome results,§ 20 on limitations and,§ 21 on generalisability.
Restructuring: Renames two section headings to reflect new ACE content	§ “recruitment” renamed to “recruitment and adaptations”§ “sample size” renamed to “sample size and operating characteristics”
Restructuring: Introduces a new subsection heading	§ “Statistical analysis plan and other trial-related documents” to accommodate item 24b
Modifies abstract *item 1b* and introduces an extension for journal and conference abstracts	
New item: Introduces one new item (on adaptation decisions made)	On “adaptation decisions made”
Modified item: Modifies one standard item	On “trial design”
Expanded text: Expands the E&E text for clarification on one item for certain ADs in particular circumstances without changes to item wording	On “Outcome”

Item numbers or section/topic referenced in *Box 5* are presented in *Table 1* and *2*.

The ACE checklist

Tables 1 and *2* are checklists for the main report and abstract, respectively. Only new and modified items are discussed in this E&E document, as well as six items that retain the CONSORT 2010 [3,4] wording but require clarification for certain ADs (*Box 5*). The item numbers are for the ACE guideline.

Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision
Harms	Important adverse events or side effects
Conclusions	General interpretation of the results
Trial registration	Registration number and name of trial register
Funding	Source of funding

Note: * Relevant to conference abstracts; “see expanded E&E text for clarification” means the wording of the item remains unchanged but we clarify additional considerations for ADs in certain circumstances in the E&E text.

Section 1. Title and abstract

CONSORT 2010 item 1b: Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts) [142]

ACE item 1b: Refer to the Adaptive designs CONSORT Extension for abstracts (see Table 2)

Explanation – A well-structured abstract summary encompassing trial design, methods, results, and conclusions is essential regardless of the type of design implemented [140,142]. This allows readers to search for relevant studies of interest and to quickly judge if the reported trial is relevant to them for further reading. Furthermore, it helps readers to make instant judgements on key benefits and risks of study interventions. *Table 2* presents minimum essential items authors should report in an AD randomised trial abstract. Authors should use this extension together with the CONSORT for journal and conference abstracts for additional details [139,140] and other relevant extensions where appropriate.

CONSORT abstract item (Trial design): Description of the trial design (for example, parallel, cluster, non-inferiority)

ACE abstract item (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

Explanation – AD randomised trials should be indexed properly to allow other researchers to easily retrieve them in literature searches. This is particularly important as trial design may influence interpretation of trial findings and the evidence synthesis approach used during meta-analyses. The MEDLINE database provides "Adaptive clinical trial" as a Medical Subject Heading (MeSH) topic to improve indexing [143]. Authors may also like to state the type of the AD, including details of adaptations as covered under the new *item 3b (Table 2)*.

Box 6. Exemplars on the use of "adaptive" in the abstract content and/or as a keyword

Example 1. Abstract (title)

"Safety and efficacy of neublastin in painful lumbosacral radiculopathy: a randomized, double-blinded, placebo-controlled phase 2 trial using Bayesian adaptive design (the SPRINT trial)." [144]

Example 2. Abstract (background)

"The drug development process can be streamlined by combining the traditionally separate stages of dose-finding (Phase IIb) and confirmation of efficacy and safety (Phase III) using an adaptive seamless design." [145]

Example 3. Abstract (aims) and keyword

"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" [102]

CONSORT/ACE abstract item (Outcome): Clearly defined primary outcome for this report

Explanation – In some AD randomised trials, the outcome used to inform adaptations (adaptation outcome) and the primary outcome of the study can differ (see *item 6* of the main checklist for details). The necessity of reporting both of these outcomes and results in the abstract depends on the stage of reporting and whether the adaptation decisions made were critical to influencing the interpretation of the final results. For example, when a trial or at least a treatment group is stopped early, based on an adaptation outcome which is not the primary outcome, it becomes essential to adequately describe both outcomes in accordance with the CONSORT 2010 statement [3,4]. Contrarily, only the description of the primary outcome in the abstract will be essential when non-terminal adaptation decisions are made (e.g., to change the sample size, update randomisation, or no dropping of treatments groups at interims) and when final (not interim) results are being reported. Furthermore, the results item (*Table 2*) should be reported consistent with the stated primary and adaptation outcome(s), where necessary.

Box 7. Exemplar on reporting outcomes in the abstract

Example 1. Bayesian RAR dose finding AD with early stopping for efficacy or futility

"The primary outcome required, first, a greater than 90% posterior probability that the most promising levocarnitine dose decreases the Sequential Organ Failure Assessment (SOFA) score at 48 hours and, second (given having met the first condition), at least a 30% predictive probability of success in reducing 28-day mortality in a subsequent traditional superiority trial to test efficacy." [146]

Example 2. Sequential-step AD

"The primary efficacy endpoint was definitive cure (absence of parasites in tissue aspirates) at 6 months. If interim analyses, based on initial cure evaluated 30 days after the start of treatment …" [147]

ACE abstract item (adaptation decisions made): Specify what trial adaptation decisions were made in light of the pre-planned decision-making criteria and observed accrued data

Explanation – A brief account of changes that were made to the trial, on what basis they were made, and when is important. The fact that the design allows for adaptations will influence interpretation of results, potentially due to operational and statistical biases. If changes should have been made, but were not, then this may further influence credibility of results. See the main checklist *item 14c* for details.

Box 8. Exemplars on reporting adaptation decisions made to the trial in the abstract

Example 1. Two-stage inferential seamless phase 2/3 AD; pre-planned adaptation decisions

"A planned interim analysis was conducted for otamixaban dose selection using a pre-specified algorithm (unknown to investigators) … The selected regimen to carry forward was an intravenous bolus of 0.080mg/kg followed by an infusion of 0.140 mg/kg per hour." [148]

Example 2. Group sequential AD; early stopping decision

"The trial was stopped early (at the third interim analysis), according to pre-specified rules, after a median follow-up of 27 months, because the boundary for an overwhelming benefit with LCZ696 had been crossed" [149]

Methods

ACE item 3b (new): Type of adaptive design used, with details of the pre-planned adaptations and the statistical information informing the adaptations

Explanation – A description of the type of AD indicates the underlying design concepts and the applicable adaptive statistical methods. Although there is an inconsistent use of nomenclature to classify ADs, together with growing related methodology [13], some currently used types of ADs are presented in *Box 3.* A clear description will also improve the indexing of AD methods and for easy identification during literature reviews.

Specification of pre-planned opportunities for adaptations and their scope is essential to preserve the integrity of AD randomised trials [22] and for regulatory assessments, regardless of whether they were triggered during the trial [14,113,114]. Details of pre-planned adaptations enable readers to assess the appropriateness of statistical methods used to evaluate operating characteristics of the AD (*item 7a*) and for performing statistical inference (*item 12b*). Unfortunately, pre-planned adaptations are commonly insufficiently described [124]. Authors are encouraged to explain the scientific rationale for choosing the considered pre-planned adaptations encapsulated under the CONSORT 2010 item "scientific background and explanation of rationale" (*item 2a*). This rationale should focus on the goals of the considered adaptations in line with the study objectives and hypotheses (*item 2b*) [112,113,124,150].

Details of pre-planned adaptations with rationale should be documented in accessible study documents for readers to be able to evaluate what was planned and unplanned (e.g., protocol, interim and final SAP or dedicated trial document). Of note, any pre-planned adaptation that modifies eligibility criteria (e.g., in population enrichment ADs [96,151]) should be clearly described.

Adaptive trials use accrued statistical information to make pre-planned adaptation(s) (*item 14c*) at interim analyses guided by pre-planned decision-making criteria and rules (*item 7b*). Reporting this statistical information for guiding adaptations and how it is gathered is paramount. Analytical derivations of statistical information guiding pre-planned adaptations using statistcial models or formulae should be described to facilitate reproducibility and interpretation of results. The use of supplementary material or references to published literature is sufficient. For example, sample size re-assessment (SSR) can be performed using different methods with or without knowledge or use of treatment arm allocation [42,45,49,152]. Around 43% (15/35) of regulatory

submissions needed further clarifications because of failure to describe how a SSR would be performed [124]. Early stopping of a trial or treatment group for futility can be evaluated based on statistical information to support lack of evidence of benefit that is derived and expressed in several ways. For example, conditional power [57,153–156], predictive power [56,154,157–159], the threshold of the treatment effect, posterior probability of the treatment effect [105], or some form of clinical utility that quantifies the balance between benefits against harms [160,161] or between patient and society perspectives on health outcomes [105].

Box 9. Exemplars on reporting item 3b elements

Example 1. Pre-planned adaptations and rationale; inferentially seamless phase 2/3 AD
"The adaptive (inferentially) seamless phase II/III design is a novel approach to drug development that combines phases II and III in a single, two-stage study. The design is adaptive in that the wider choice of doses included in stage 1 is narrowed down to the dose(s) of interest to be evaluated in stage 2. The trial is a seamless experience for both investigators and patients in that there is no interruption of ongoing study treatment between the two phases. Combining the dose-finding and confirmatory phases of development into a single, uninterrupted study has the advantages of speed, efficiency and flexibility [15,17] … The primary aim of stage 1 of the study was to determine the risk-benefit of four doses of indacaterol (based on efficacy and safety results in a pre-planned interim analysis) in order to select two doses to carry forward into the second stage of the study" [145]
Example 2. Analytical derivation of statistical information to guide adaptations; population enrichment AD with SSR
Mehta et al. [104] detail formulae used to calculate the conditional power to guide modification of the sample size or to enrich the patient population at an interim analysis for both cutaneous and non-cutaneous patients (full population) and only cutaneous patients (subpopulation) in the supplementary material. In addition, the authors detail formulae used to derive associated conditional powers and p-values used for decision-making to claim evidence of benefit both at the interim and final analysis (linked to *item 12b*).
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 [69,71]. 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." [103]
Example 4. Type of AD; pre-planned adaptations and rationale; Bayesian adaptive-enrichment AD allowing for enrichment and early stopping for futility or efficacy
"The DAWN trial was a multicenter, prospective, randomized, open-label trial with a Bayesian adaptive–enrichment design and with blinded assessment of endpoints [12]. The adaptive trial design allowed for a sample size ranging from 150 to 500 patients. During interim analyses, the decision to stop or continue enrolment was based on a pre-specified calculation of the probability that thrombectomy plus standard care would be superior to standard care alone with respect to the first primary endpoint (described in the paper). The enrichment trial design gave us the flexibility to identify whether the benefit of the trial intervention was restricted to a subgroup of patients with relatively small infarct volumes at baseline. The interim analyses, which included patients with available follow-up data at the time of the analysis, were pre-specified to test for the futility, enrichment, and success of the trial." [105] See supplementary appendix via journal website (from page 39) for details.
Example 5. Rationale; type of AD and pre-planned adaptations; information to inform adaptations; information-based GSD
"Because little was known about the variability of LVMI changes in CKD during the planning stage, we prospectively implemented an information-based (group sequential) adaptive design that allowed sample size re-estimation when 50% of the data were collected [51,162]" [163]. Pritchett et al. [51] provide details of the pre-planned adaptations and statistical information used to inform SSR and efficacy early stopping.
Example 6. Pre-planned adaptation and information for SSR
"To reassess the sample size estimate, the protocol specified that a treatment-blinded interim assessment of the standard deviation (SD) about the primary endpoint (change from baseline in total exercise treadmill test duration at trough) would be performed when 231 or one half of the planned completed study patients had been randomized and followed up for 12 weeks. The recalculation of sample size, using only blinded data, was adjusted based on the estimated SD of the primary efficacy parameter (exercise duration at trough) from the aggregate data … [164–166]." [39]

CONSORT 2010 item 3b: Important changes to the design or methods after trial commencement (such as eligibility criteria), with reasons

ACE item 3c (modification, renumbered): Important changes to the design or methods after trial commencement (such as eligibility criteria) outside the scope of the pre-planned adaptive design features, with reasons

Explanation – Unplanned changes to certain aspects of design or methods in response to unexpected circumstances that occur during the trial are common and will need to be reported in AD randomised trials, as in fixed designed trials. This may include deviations from pre-planned adaptations and decision rules [15,63], as well as changes to timing and frequency of interim analyses. Traditionally, unplanned changes with explanation have been documented as protocol amendments and reported as discussed in the CONSORT 2010 statement [3,4]. Unplanned changes, depending on what they are and why they were made may introduce bias and compromise trial credibility. Some unplanned changes may render the planned adaptive statistical methods invalid or may complicate interpretation of results [22]. It is therefore essential for authors to detail important changes that occurred outside the scope of the pre-planned adaptations and to explain why deviations from the planned adaptations were necessary. Furthermore, it should be clarified whether unplanned changes were made following access to key trial information (e.g., interim data seen by treatment group or interim results). Such information will help readers assess potential sources of bias and implications for interpretation of results. For ADs, it is essential to distinguish unplanned changes from pre-planned adaptations (*item 3b*) [167].

Box 10. Exemplars on reporting item 3c elements

Example. Inferentially seamless phase 2/3 (5-arm 2-stage) AD allowing for regimen selection, SSR and futility early stopping

Although this should ideally have been referenced in the main report, Léauté-Labrèze et al. [103] (on pages 17-18 of supplementary material) summarise important changes to the trial design including an explanation and discussion of implications. These changes include a reduction in the number of patients assigned to the placebo across stages – randomisation was changed from 1:1:1:1:1 to 2:2:2:2:1 (each of the 4 propranolol regimens: placebo) for stage 1 and from 1:1 to 2:1 for stage 2 in favour of the selected regimen; revised complete or nearly complete resolution success rates for certain treatment regimens – as a result, total sample size was revised to 450 (excluding possible SSR); and a slight increase in the number of patients (from 175 to 180) to be recruited for the interim analysis.

Section 6. Outcomes

CONSORT 2010 item 6a: Completely define pre-specified primary and secondary outcome measures, including how and when they were assessed

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

Comment – Authors should also refer to the CONSORT 2010 statement [3,4] for the original text when applying this item.

Explanation – It is paramount to provide a detailed description of pre-specified outcomes used to assess clinical objectives including how and when they were assessed. For operational feasibility, ADs often use outcomes that can be observed quickly and easily to inform pre-planned adaptations (adaptation outcomes). Thus, in some situations, adaptations may be based on early observed outcome(s) [168] that are believed to be informative for the primary outcome even though different from the primary outcome. The adaptation outcome (such as a surrogate, biomarker, or an intermediate outcome) together with the primary outcome influences the adaptation process, operating characteristics of the AD, and interpretation and trustworthiness of trial results. Despite many potential advantages of using early observed outcomes to adapt a trial, they pose additional risks of making misleading inferences if they are unreliable [169]. For example, a potentially beneficial treatment could be wrongly discarded, an ineffective treatment incorrectly declared effective or wrongly carried forward for further testing, or the randomisation updated based on unreliable information.

Authors should therefore clearly describe adaptation outcomes similar to the description of pre-specified primary and secondary outcomes in the CONSORT 2010 statement [3,4]. Authors are encouraged to provide a clinical rationale supporting the use of adaptation outcome that is different to the primary outcome in order to aid the clinical interpretation of results. For example, evidence supporting that the adaptation outcome can provide reliable information on the primary outcome will suffice.

Box 11. Exemplars on reporting item 6a elements

Example 1. SSR; description of the adaptation and primary outcomes

"The primary endpoint is a composite of survival free of debilitating stroke (modified Rankin score >3) or the need for a pump exchange. The short-term endpoint will be assessed at 6 months and the long-term endpoint at 24 months (primary). Patients who are urgently transplanted due to a device complication before a pre-specified endpoint will be considered study failures. All other transplants or device explants due to myocardial recovery that occur before a pre-specified endpoint will be considered study successes ... The adaptation was based on interim short-term outcome rates." [170]

Example 2. Seamless phase 2/3 Bayesian AD with treatment selection; details of adaptation outcomes

"Four efficacy and safety measures were considered important for dose selection based on early phase dulaglutide data: HbA1c, weight, pulse rate and diastolic blood pressure (DBP) [171]. These measures were used to define criteria for dose selection. The selected dulaglutide dose(s) had to have a mean change of ≤+5 beats per minute (bpm) for PR and ≤+2 mmHg for DBP relative to placebo at 26 weeks. In addition, if a dose was weight neutral versus placebo, it had to show HbA1c reduction ≥1.0% and/or be superior to sitagliptin at 52 weeks. If a dose reduced weight relative to placebo ≥ 2.5 kg, then non-inferiority to sitagliptin would be acceptable. A clinical utility index was incorporated in the algorithm to facilitate adaptive randomization and dose selection [160,172] based on the same parameters used to define dose-selection criteria described above (not shown here)." [102]

Example 3. Seamless phase 2/3 AD with treatment selection; details of adaptation outcomes

"For the dose selection, the joint primary efficacy outcomes were the trough FEV_1 on Day 15 (mean of measurements at 23 h 10 min and 23 h 45 min after the morning dose on Day 14) and standardized (average) FEV_1 area under the curve (AUC) between 1 and 4 h after the morning dose on Day 14 (FEV_1AUC_{1-4h}), for the treatment comparisons detailed below (not shown here)." [145]

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." [173]

CONSORT 2010 item 6b: Any changes to trial outcomes after the trial commenced, with reasons

ACE item 6b (modification): Any unplanned changes to trial outcomes after the trial commenced, with reasons

Comment – Authors may wish to cross-reference the CONSORT 2010 statement [3,4] for background details.

Explanation – Outcome reporting bias occurs when the selection of outcomes to report is influenced by the nature and direction of results. The prevalence of outcome reporting bias in medical research is well documented: discrepancies between pre-specified outcomes in protocols or registries and those published in reports [12,174–177]; outcomes that portray favourable beneficial effects of treatments and safety profiles being more likely to be reported [178]; some pre-specified primary or secondary outcomes modified or switched after trial commencement [179]. Only 13% (9/67) of trials published had all primary and secondary outcomes completely reported consistent with the protocol or registry. Changes to trial outcomes may also include changes to how outcomes were assessed or measured, when they were assessed, or order of importance to address objectives [177].

Sometimes when planning trials, there is huge uncertainty around the magnitude of treatment effects on potential outcomes viewed acceptable as primary endpoints [37,177]. As a result, although uncommon, a pre-planned adaptation could include the choice of the primary endpoints or hypotheses for assessing the benefit-risk ratio. In such circumstances, the adaptive strategy should be clearly described as a pre-planned adaptation (*item 3b*). Authors should clearly report any additional changes to outcomes outside the scope of the pre-specified adaptations including an explanation of why such changes occurred in line with the CONSORT 2010 statement. This will enable readers to distinguish pre-planned trial adaptations of outcomes from unplanned changes, thereby allowing them to judge outcome reporting bias.

Box 12. Exemplar on reporting item 6b

Example. Bayesian adaptive-enrichment AD; unplanned change from a secondary to a co-primary outcome, rationale, and when it happened
"The second primary endpoint was the rate of functional independence (defined as a score of 0, 1, or 2 on the modified Rankin scale) at 90 days. This endpoint was changed from a secondary endpoint to a co-primary endpoint at the request of the Food and Drug Administration at 30 months after the start of the trial, when the trial was still blinded." [105]

Section 7. Sample size and operating characteristics

CONSORT 2010 item 7a: How the sample size was determined

ACE item 7a (modification): How sample size and operating characteristics were determined

Comments – This section heading was modified to reflect additional operating characteristics that may be required for some ADs in addition to the sample size. *Items 3b*, *7a*, *7b* and *12b* are connected so they should be cross-referenced when reporting.

Explanation – Operating characteristics, which relate to statistical behaviour of a design, should be tailored to address trial objectives and hypotheses, factoring in logistical, ethical and clinical considerations. These may encompass the maximum sample size, expected sample sizes under certain scenarios, probabilities of identifying beneficial treatments if they exist, and probabilities of making false positive claims of evidence [180,181]. Specifically, the predetermined sample size for ADs is influenced, among other things, by:

1. type and scope of adaptations considered (*item 3b*);
2. decision-making criteria used to inform adaptations (*item 7b*);
3. criteria for claiming overall evidence (e.g., based on the probability of the treatment effect being above a certain value, targeted treatment effect of interest, and threshold for statistical significance [182,183])
4. timing and frequency of the adaptations (*item 7b*);
5. type of primary outcome(s) (*item 6a*) and nuisance parameters (e.g., outcome variance);
6. method for claiming evidence on multiple key hypotheses (part of *item 12b*);
7. desired operating characteristics (see *Box 2*), such as statistical power and an acceptable level of making a false positive claim of benefit;
8. adaptive statistical methods used for analysis (*item 12b*);
9. statistical framework (frequentist or Bayesian) used to design and analyse the trial.

Information that guided estimation of sample size(s) including operating characteristics of the considered AD should be described sufficiently to enable readers to reproduce the sample size calculation. The assumptions made concerning design parameters should be clearly stated and supported with evidence

if possible. Any constraints imposed (e.g., due to limited trial population) should be stated. It is good scientific practice to reference the statistical tools used (e.g., statistical software, program, or code) and to describe the use of statistical simulations when relevant (see *item 24b* discussion).

In a situation where changing the sample size is a pre-planned adaptation (*item 3b*), report the initial sample sizes (at interims analyses before the expected change in sample size) and the maximum allowable sample size per group and in total if applicable. The planned sample sizes (or expected number of events for time-to-event data) at each interim analysis and final analysis should be reported by treatment group and overall. The timing of interim analyses can be specified as a fraction of information gathered rather than sample size.

Box 13. Exemplars on reporting item 7a elements

Example 1. MAMS AD; assumptions and adaptive methods; approach for claiming evidence or informing adaptations; statistical program
"The primary response (outcome) from each patient is the difference between the baseline HOMA-IR score and their HOMA-IR score at 24 weeks. The sample size calculation is based on a one-sided type I error of 5% and a power of 90%. If there is no difference between the mean response on any treatment and that on control, then a probability of 0.05 is set for the risk of erroneously ending the study with a recommendation that any treatment be tested further. For the power, we adopt a generalisation of this power requirement to multiple active treatments due to Dunnett [184]. Effect sizes are specified as the percentage chance of a patient on active treatment achieving a greater reduction in HOMA-IR score than a patient on control as this specification does not require knowledge of the common SD, σ. The requirement is that, if a patient on the best active dose has a 65% chance of a better response than a patient on control, while patients on the other two active treatments have a 55% chance of showing a better response than a patient on control, then the best active dose should be recommended for further testing with 90% probability. A 55% chance of achieving a better response on active dose relative to control corresponds to a reduction in mean HOMA-IR score of about a sixth of an SD (0.178σ), while the clinically relevant effect of 65% corresponds to a reduction of about half an SD (0.545σ). The critical values for recommending that a treatment is taken to further testing at the interim and final analyses (2.782 and 2.086) have been chosen to guarantee these properties using a method described by Magirr et al [185], generalising the approach of Whitehead and Jaki [186]. The maximum sample size of this study is 336 evaluable patients (84 per arm), although the use of the interim analysis may change the required sample size. The study will recruit additional patients to account for an anticipated 10% dropout rate (giving a total sample size of 370). An interim analysis will take place once the primary endpoint is available for at least 42 patients on each arm (i.e., total of 168, half of the planned maximum of 336 patients). Sample size calculation was performed using the MAMS package in R [187]." [58]
Example 2. 3-arm 2-stage AD with dose selection; group sequential approach; assumptions; adaptation decision-making criteria; stage 1 and 2 sample sizes; use of simulations
"Sample size calculations are based on the primary efficacy variable (composite of all-cause death or new MI through day 7), with the following assumptions: an event rate in the control group of 5.0%, based on event rates from the phase II study (24); a relative risk reduction (RRR) of 25%; a binomial 1-sided (α=0.025) superiority test for the comparison of 2 proportions with 88% power; and a 2-stage adaptive design with one interim analysis at the end of stage 1 data (35% information fraction) to select 1 otamixaban dose for continuation of the study at stage 2. Selection of the dose for continuation was based on the composite end point of all-cause death, Myocardial Infarction (MI), thrombotic complication, and the composite of Thrombosis in Myocardial Infarction (TIMI) major bleeding through day 7, with an assumed probability for selecting the "best" dose according to the primary endpoint (r = 0.6), a group sequential approach with futility boundary of relative risk of otamixaban versus UFH plus eptifibatide ≥1.0, and efficacy boundary based on agamma (-10) α spending function [188]. Based on the above assumptions, simulations (part of *item 24b*, see supplementary material) showed that 13,220 patients (a total of 5625 per group for the 2 remaining arms for the final analysis) are needed for this study." [189] See *Figure 1*.

Fig 1 legend: Adapted from Steg et al. [197].

CONSORT 2010 item 7b: When applicable, explanation of any interim analyses and stopping guidelines

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

Comments – This item is a replacement so when reporting, the CONSORT 2010 [3] *item 7b* content should be ignored. *Items 7b and 8b* overlap but we intentionally reserved *item 8b* specifically to enhance complete reporting of ADs with randomisation updates as a pre-planned adaptation. Reporting of these items is also connected to *items 3b* and *12b*.

Explanation – Transparency and complete reporting of pre-planned decision-making criteria (*Box 2*) and how overall evidence is claimed are essential as they influence operating characteristics of the AD, credibility of the trial, and clinical interpretation of findings [22,32,190].

A key feature of an AD is that interim decisions about the course of the trial are informed by observed interim data (element of *item 3b*) at one or more interim analyses guided by decision rules describing how and when the proposed adaptations will be activated (pre-planned adaptive decision-making criteria). Decision rules, as defined in *Box 2*, may include, but are not limited to, rules for making adaptations described in *Box 3*. Decision rules are often constructed with input of key stakeholders (e.g., clinical investigators, statisticians, patient groups, health economists, and regulators) [191]. For example, statistical methods for formulating early stopping decision rules of a trial or treatment group(s) exist [52,53,192–195].

Decision boundaries (e.g., stopping boundaries), pre-specified limits or parameters used to determine adaptations to be made, and criteria for claiming overall evidence of benefit and/or harm (at interim or final analysis) should be clearly stated. These are influenced by statistical information used to inform adaptations (*item 3b*). Decision trees or algorithms can aid the representation of complex adaptive decision-making criteria

Allowing for trial adaptations too early in a trial with inadequate information severely undermines robustness of adaptive decision-making criteria and trustworthiness of trial results [196,197]. Furthermore, methods and results can only be reproducible when timing and frequency of interim analyses are adequately described. Therefore, authors should detail when and how often the interim analyses were planned to be implemented. The planned timing can be described in terms of information such as interim sample size or number of events relative to the maximum sample size or maximum number of events, respectively. For example, in circumstances when the pre-planned and actual timing or/and frequency of the interim analyses differ, reports should clearly state what actually happened (*item 3c*).

Clarification should be made on whether decision rules were binding or non-binding to help assess implications in the case when they are overruled or ignored. For example, when a binding futility boundary is overruled and a trial is continued, this would lead to a type I error inflation. Thus, non-binding decision rules are those that can be overruled without having a negative effect on the control of the type I error rate. Use of non-binding futility boundaries is often advised [56].

Box 14. Exemplars on reporting item 7b elements

Example 1. 2-arm 2-stage AD with options for early stopping for futility or superiority and to increase the sample size; binding stopping rules

"To calculate the number of patients needed to meet the primary end-point, we expected a 3-year overall survival rate of 25% in the group assigned to preoperative chemotherapy (arm A) (based on two previous trials [198,199]). In comparison, an increase of 10% (up to 35%) was anticipated by preoperative CRT. Using the log-rank test (one-sided at this point) at a significance level of 5%, we calculated to include 197 patients per group to ensure a power of 80%. In the first stage of the planned two-stage adaptive design [200], the study was planned to be continued on the basis of a new calculation of patients needed if the comparison of patient groups will be $0.0233< p_1< 0.5$. Otherwise, the study may be closed for superiority ($p_1< 0.0233$) or shall be closed for futility ($p_1 \geq 0.5$). There was no maximum sample size cap and stopping rules were binding." [201]. p_1 and p_2 are p-values derived from independent stage 1 and stage 2 data, respectively. Evidence of benefit will be claimed if the overall two-stage p-value derived from p_1 and p_2 is ≤ 0.05.

Example 2. Timing and frequency of interim analyses; planned stopping boundaries for superiority and futility

Table 3. Stopping boundaries

Interim Analysis	Number of primary outcome events (information fraction)	Stopping boundaries			
		Superiority		Futility	
		Hazard ratio	P-value	Hazard ratio	P-value
1	800 (50%)	<0.768	<0.0002	>0.979	>0.758
2	1200 (75%)	<0.806	<0.0002	>0.931	>0.216
Final	1600 (100%)	<0.906	<0.0500		

Adapted from Pocock et al. [202]; primary outcome events are cardiovascular deaths, myocardial infarction, or ischemic stroke.

Example 3. Planned timing and frequency of interim analysis; pre-specified dose selection rules for an inferentially seamless phase 2/3 (7-arm 2-stage) AD

"The interim analysis was pre-planned for when at least 110 patients per group (770 total) had completed at least 2 weeks of treatment. The dose selection guidelines were based on efficacy and safety. The mean effect of each indacaterol dose versus placebo was judged against pre-set efficacy reference criteria for trough FEV_1 and FEV_1AUC_{1-4h}. For trough FEV_1, the reference efficacy criterion was the highest value of: (a) the difference between tiotropium and placebo, (b) the difference between formoterol and placebo, or (c) 120 mL (regarded as the minimum clinically important difference). For standardized FEV_1AUC_{1-4h}, the reference efficacy criterion was the highest value of: (a) the difference between tiotropium and placebo or (b) the difference between formoterol and placebo. If more than one indacaterol dose exceeded both the efficacy criteria, the lowest effective dose plus the next higher dose were to be selected. Data on peak FEV_1, % change in FEV_1, and FVC were also supplied to the DMC for possible consideration, but these measures were not part of the formal dose selection process and are not presented here. The DMC also took into consideration any safety signals observed in any treatment arm." [145]

Example 4. Timing and frequency of interim analyses; decision-making criteria for population enrichment and sample size increase

"Cohort 1 will enrol a total of 120 patients and followed them until 60 PFS events are obtained. At an interim analysis based on the first 40 PFS events, an independent data monitoring committee will compare the conditional power for the full population (CP_F) and the conditional power for the cutaneous subpopulation (CP_S). The formulae for these conditional powers are given in the supplementary appendix (*part of item 3b, example 2, Box 9*). (a) If CP_F <0.3 and CP_S <0.5, the results are in the unfavourable zone; the trial will enrol 70 patients to cohort 2 and follow them until 35 PFS events are obtained (then test effect in the full population). (b) If CP_F <0.3 and CP_S >0.5, the results are in the enrichment zone; the trial will enrol 160 patients with cutaneous disease (subpopulation) to cohort 2 and follow them until 110 PFS events have been obtained from the combined patients in both cohorts with cutaneous disease only (then test effect only in the cutaneous subpopulation). (c) If $0.3 \leq CP_F \leq 0.95$, the results are in the promising zone (so increase sample size); the trial will enrol 220 patients (full population) to cohort 2 and follow them up until 110 PFS events are obtained (then test effect in the full population). (d) If CP_F >0.95, the results are in the favourable zone; the trial will enrol 70 patients to cohort 2 and follow them until 35 PFS events are obtained (then test effect in full population)" [104]. See *Figure 2* of Mehta et al. [104] for a decision-making tree.

Example 5. Bayesian GSD with futility early stopping; frequency and timing of interim analyses; adaptation decision-making criteria; criteria for claiming treatment benefit

"We adopted a group-sequential Bayesian design [194] with three stages, of 40 patients each (in total), and two interim analyses after 40 and 80 randomised participants, and a final analysis after a maximum of 120 randomised participants. We decided that the trial should be stopped early if there is a high (posterior) probability (90% or greater) (*item 3b* details) that the 90-day survival odds ratio (OR) falls below 1 (i.e. REBOA is harmful) at the first or second interim analysis. REBOA will be declared "successful" if the probability that the 90-day survival OR exceeds 1 at the final analysis is 95% or greater." [203]

Example 6. *Sample size cap following SSR and treatment selection*

"… the IDMC may recommend to increase the sample size at the interim analysis if there is less than 80% conditional power (CP) (*item 3b* details) to demonstrate superiority of the selected regimen (or at least one of the selected regimens in the case where two regimens are chosen) over placebo. The maximum permitted sample size increase will be fixed at 100 additional patients per selected propranolol regimen and 50 additional patients on the placebo arm" [103]. Extracted from supplementary material.

Example 7. 2-stage AD; use of non-binding futility boundary

"It has to be noted that, to protect the global type 1 error in case the decision was taken to overrule the futility rule, nonbinding boundaries would be used adding a very conservative boundary for efficacy. Overwhelming efficacy would be assessed on the adjudicated primary efficacy endpoint using the gamma (-10) alpha spending function [188] , in comparing the observed one-sided p-value with 0.00004 during this interim analysis. With such a conservative alpha spending function, the global alpha level of the study would be maintained at 0.05 (2-sided)." [148]. Extracted from supplementary material.

Section 8. Randomisation (Sequence generation)

CONSORT 2010 item 8b: Type of randomisation; details of any restriction (such as blocking and block size)

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

Comments – In applying this item, the reporting of randomisation aspects before activation of trial adaptations must adhere to CONSORT 2010 *items 8a* and *8b*. This E&E document only addresses additional randomisation aspects that are essential when reporting any AD where the randomisation allocation changes. Note that the contents of extension *items 7b* and *8b* overlap.

Explanation – In AD randomised trials, the allocation ratio(s) may remain fixed throughout or change during the trial as a consequence of pre-planned adaptations (e.g., when modifying randomisation to favour treatments more likely to show benefits, after treatment selection, or introduction of a new arm to an ongoing trial) [76]. Unplanned changes may also change allocation ratios (e.g., after early stopping of a treatment arm due to unforeseeable harms).

This reporting item is particularly important for response-adaptive randomisation (RAR) ADs as several factors influence their efficiency and operating characteristics, which in turn influence the trustworthiness of results and necessitate adequate reporting [13,204–207]. For RAR ADs, authors should therefore detail the pre-planned:

1. burn-in period before activating randomisation updates including period when the control group allocation ratio was fixed. Use of RAR with inadequate or without burn-in period has raised debate on credibility of some trials [206,207];
2. type of randomisation method with allocation ratios per group during the burn-in period as detailed in the updated standard CONSORT *item 8b*;
3. method or algorithm used to adapt or modify the randomisation allocations after the burn-in period;
4. information used to inform the adaptive randomisation algorithm and how it was derived (*item 3b*). Specifically, when a Bayesian RAR is used, we encourage authors to provide details of statistical models and rationale for the prior distribution chosen;
5. frequency of updating the allocation ratio (e.g., after accrual of a certain number of participants with outcome data or defined regular time period) and;
6. adaptive decision-making criteria to declare early evidence in favour or against certain treatment arms (part of *item 7b*).

In addition, any envisaged changes to the allocation ratio as a consequence of other trial adaptations (e.g., early stopping of an arm or addition of a new arm) should be stated.

Box 15. Exemplars on reporting item 8b elements
Example 1. Pre-planned changes to allocation ratios as a consequence of treatment selection or/and sample size increase
“All new patients recruited after the conclusions of the interim analysis are made, will be randomised in a (2:) 2: 1 ratio to the selected regimen(s) of propranolol or placebo until a total of (100:)100: 50 patients (or more in the case where a sample size increase is recommended) have been randomised over the two stages of the study.” [103] Extracted from supplementary material. (2:) and (100:) are only applicable if the second best regimen is selected at stage 1.
Example 2. Bayesian RAR; Pre-planned algorithm to update allocation ratios; frequency of updates (after every participant);no burn-in period; period of a fixed control allocation ratio; information that informed adaptation; decision-making criteria for dropping treatments (part of item 7b)
See Additional file 1 as extracted from Giles et al. [73].
Example 2. Bayesian RAR; burn-in period; fixed control allocation ratio; details of adaptive randomisation including additional adaptations and decision-making criteria (part of item 7b); derivation of statistical quantities; details of Bayesian models and prior distribution with rationale
“ … eligible patients were randomized on day 1 to treatment with placebo or neublastin 50, 150, 400, 800, or 1200 mg/kg, administered by intravenous injection on days 1, 3, and 5. The first 35 patients were randomized in a 2:1:1:1:1:1 ratio to placebo and each of the 5 active doses (randomisation method required) (i.e., 10 patients in the placebo group and 5 for each dose of active treatment). Subsequently, 2 of every 7 enrolled patients were assigned to placebo. Interim data evaluations of pain (AGPI) and pruritus questionnaire data (proportion of patients who reported “the itch is severe enough to cause major problems for me” on an Itch Impact Questionnaire) were used to update the allocation probability according to a Bayesian algorithm for adaptive allocation and to assess efficacy and futility criteria for early stopping of enrolment (Fig. 1, not shown here). Interim evaluations and updates to the allocation probabilities were performed weekly. (. Enrolment was to be stopped early after ≥50 patients had been followed for 4 weeks if either the efficacy criterion (>80% probability that the maximum utility dose reduces the pain score by ≥1.5 points more than the placebo) or the futility criterion (<45% probability that the maximum utility dose reduces pain more than the placebo) was met.” [144] Details of statistical models used including computation of posterior quantities; prior distribution with rationale; generation of the utility function; and weighting of randomisation probabilities are accessible a weblink provided (https://links.lww.com/PAIN/A433).

Section 11. Blinding

ACE item 11c (new): Measures to safeguard the confidentiality of interim information and minimise potential operational bias during the trial

Explanation – Preventing or minimising bias is central for robust evaluation of the beneficial and harmful effects of interventions. Analysis of accumulating trial data brings challenges regarding how knowledge or leakage of information, or mere speculation about interim treatment effects, may influence behaviour of key stakeholders involved in the conduct of the trial [22,128,208]. Such behavioural changes may include differential clinical management; reporting of harmful effects; clinical assessment of outcomes; and decision-making to favour one treatment group over the other. Inconsistencies in trial conduct before and after adaptations have wide implications that may affect trial validity and integrity [22]. For example, use of statistical methods that combine data across stages may become questionable or may make overall results uninterpretable. AD randomised trials whose integrity was severely compromised by disclosure of interim results resulted in regulators questioning the credibility of conclusions [209,210]. Most AD randomised trials, 76% (52/68) [117] and 60% (151/251) [125], did not disclose methods to minimise potential operational bias during interim analyses. The seriousness of this potential risk will depend on various trial characteristics and the purpose of having disclosure is to enable readers to judge the risk of potential sources of bias, and thus judge how trustworthy they can assume results to be.

Literature covers processes and procedures which could be considered by researchers to preserve confidentiality of interim results to minimise potential operational bias [46,150,211]. There is no universal approach that suits every situation due to factors such as feasibility; nature of the trial; and available resources and infrastructure. Some authors discuss roles and activities of independent committees in adaptive decision-making processes and control mechanisms for limiting access to interim information [211–213].

Description of the process and procedures put in place to minimise the potential introduction of operational bias related to interim analyses and decision-making to inform adaptations is essential [22,131,211]. Specifically, authors should give consideration to:

1. who recommended or made adaptation decisions? The roles of the sponsor or funder, clinical investigators, and trial monitoring committees (e.g., independent data monitoring committee or dedicated committee for adaptation) in the decision-making process should be clearly stated
2. who had access to interim data and performed interim analyses
3. safeguards which were in place to maintain confidentiality (e.g., how the interim results were communicated and to whom and when)

Box 16. Exemplars on reporting item 11c elements

Example 1. Inferentially seamless phase 2/3 AD
"The interim analysis was carried out by an independent statistician (from ClinResearch GmbH, Köln, Germany), who was the only person outside the Data Monitoring Committee (DMC) with access to the semi-blinded randomization (sic) codes (treatment groups identified by letters A to G). This statistician functioned independently of the investigators, the sponsor's clinical trial team members and the team that produced statistical programming for the interim analysis (DATAMAP GmbH, Freiburg, Germany). The independent statistician was responsible for all analyses of efficacy and safety data for the interim analysis. The DMC was given semi-blinded results with treatment groups identified by the letters A to G, with separate decodes sealed in an envelope to be opened for decision-making. The personnel involved in the continuing clinical study were told which two doses had been selected, but study blinding remained in place and the results of the interim analysis were not communicated. No information on the effects of the indacaterol doses (including the two selected) was communicated outside the DMC." [145]
Example 2. Bayesian inferentially seamless phase 2/3 AD with RAR
"An independent Data Monitoring Committee (DMC) external to Lilly provided oversight of the implementation of the adaptive algorithm and monitored study safety. The DMC fulfilled this role during the dose-finding portion, and continued monitoring after dose selection until an interim database lock at 52 weeks, at which time the study was unblinded to assess the primary objectives. Sites and patients continued to be blinded to the treatment allocation until the completion of the study. The DMC was not allowed to intervene with the design operations. A Lilly Internal Review Committee (IRC), independent of the study team, would meet if the DMC recommended the study to be modified. The role of the IRC was to make the final decision regarding the DMC's recommendation. The external Statistical Analysis Center (SAC) performed all interim data analyses for the DMC, evaluated the decision rules and provided the randomization updates for the adaptive algorithm. The DMC chair and the lead SAC statistician reviewed these (interim) reports and were tasked to convene an unscheduled DMC meeting if an issue was identified with the algorithm or the decision point was triggered" [102]
Example 3. Inferentially seamless phase 2/3 AD with treatment selection, SSR, and nonbinding futility stopping
"Following the interim analysis of the data and the review of initial study hypotheses, the committee (IDMC) chairman will recommend in writing to the sponsor whether none, one or two regimen(s) of propranolol is (are) considered to be the 'best' (the most efficacious out of all regimens with a good safety profile) for further study in stage two of the design. The second 'best' regimen will only be chosen for further study along with the 'best' regimen if the first stage of the study suggests that recruitment in the second stage will be too compromised by the fact that 1 in 3 patients are assigned to placebo. The IDMC will not reveal the exact sample size increase in the recommendation letter in order to avoid potential sources of bias (only the independent statistician, the randomisation team and the IP suppliers will be informed of the actual sample size increase). Any safety concerns will also be raised in the IDMC recommendation letter. The chairman will ensure that the recommendations do not unnecessarily unblind the study. In the case where the sponsor decides to continue the study, the independent statistician will communicate to the randomisation team which regimen(s) is (are) to be carried forward." [103] Extracted from supplementary material.

Section 12. Statistical methods

CONSORT 2010 item 12a: Statistical methods used to compare groups for primary and secondary outcomes

ACE item 12a (modification): Statistical methods used to compare groups for primary and secondary outcomes, and any other outcomes used to make pre-planned adaptations

Comment – This item should be applied with reference to the detailed discussion in the CONSORT 2010 statement [3,4].

Explanation – The CONSORT 2010 statement [3,4] addresses the importance of detailing statistical methods to analyse primary and secondary outcomes at the end of the trial. This ACE modified item extends this to require similar description to be made of statistical methods used for interim analyses. Furthermore, statistical methods used to analyse any other adaptation outcomes (*item 6*) should be detailed to enhance reproducibility of the adaptation process and results. Authors should focus on complete description of statistical models and aspects of the estimand of interest [214,215] consistent with stated objectives and hypotheses (*item 2b*) and pre-planned adaptations (*item 3b*).

For Bayesian ADs, *item 12b* (paragraph 6) describes similar information that should be reported for Bayesian methods.

Box 17. Exemplars on reporting item 12a elements
Example 1. Frequentist AD
Authors are referred to the CONSORT 2010 statement [3,4] for examples.
Example 2. 2-stage Bayesian biomarker-based AD with RAR
In a methods paper, Gu et al. [216] detail Bayesian logistic regression models for evaluating treatment and marker effects at the end of stage 1 and 2 using non-informative normal priors during RAR and futility early stopping decisions. Strategies for variable selection and model building at the end of stage 1 to identify further important biomarkers for use in RAR of stage 2 patients are described (part of *item 3b*), including a shrinkage prior used for biomarker selection with rationale.

ACE item 12b (new): For the implemented adaptive design features, statistical methods used to estimate treatment effects for key endpoints and to make inferences

Comment – Note that *items 7a* and *12b* are connected.

Explanation – A goal of every trial is to provide reliable estimates of the treatment effect for assessing benefits and risks to reach correct conclusions. Several statistical issues may arise when using an AD depending on its type and the scope of adaptations, the adaptive decision-making criteria and whether frequentist or Bayesian methods are used to design and analyse the trial [22]. Conventional estimates of treatment effect based on fixed design methods may be unreliable when applied to ADs (e.g., may exaggerate the patient benefit) [101,217–221]. Precision around the estimated treatment effects may be incorrect (e.g., the width of confidence intervals may be incorrect). Other methods available to summarise the level of evidence in hypothesis testing (e.g., p-values) may give different answers. Some factors and conditions that influence the magnitude of estimation bias have been investigated and there are circumstances when it may not be of concern [217,222–226]. Cameron et al. [227] discuss methodological challenges in performing network meta-analysis when combining evidence from randomised trials with ADs and fixed designs. Statistical methods for estimating the treatment effect and its precision exist for some ADs [71,228–237] and implementation tools are being developed [86,238–240]. However, these methods are rarely used or reported and the implications are unclear [50,217,241]. Debate and research on inference for some ADs with complex adaptations is ongoing.

In addition to statistical methods for comparing outcomes between groups (*item 12a*), we specifically encourage authors to clearly describe statistical methods used to estimate measures of treatment effects with associated uncertainty (e.g., confidence or credible intervals) and p-value (when appropriate); referencing relevant literature is sufficient. When conventional or naïve estimates derived from fixed design methods are used, it should be clearly stated. In situations where statistical simulations were used to either explore the extent of bias in estimation of the treatment effects or operating characteristics, it is good practice to mention this and provide supporting evidence (*item 24c*). For example, some authors quantified bias in the estimation of treatment effects in a specific trial context through simulations [189,242].

ADs tend to increase the risk of making misleading or unjustified claims of treatments effects if traditional methods that ignore trial adaptations are used. In general, this arises when selecting one or more hypothesis test results from a possible list in order to claim evidence of the desired conclusion. For instance, the risks may increase by testing same hypothesis several times (e.g., at interim and final analyses), hypothesis testing of multiple treatment comparisons, selecting an appropriate population from multiple target populations, adapting key outcomes, or a combination of these [22]. A variety of adaptive statistical methods exist for controlling specific operating characteristics of the design (e.g., type I error rate, power) depending on the source of inflation of the false positive error rate [52,62,86,98,200,243–248].

Authors should therefore state operating characteristics of the design that have been controlled and details of statistical methods used. The need for controlling a specific type of operating characteristic (e.g., pairwise or familywise type I error rate) is context dependent (e.g., regulatory considerations, objectives and setting) so clarification is encouraged to help interpretation. How evidence of benefit and/or risk is claimed (part of *item 7a*) and hypotheses

being tested (*item 2b*) should be clear. In situations where statistical simulations were used, we encourage authors to provide a report, where possible (*item 24b*).

When data or statistical tests across independent stages are combined to make statistical inference, authors should clearly describe the combination test method (e.g. Fishers combination method, inverse normal method or conditional error function) [200,246,247,249,250] and weights used for each stage (when not obvious). This information is important because different methods and weights may produce results that lead to different conclusions. Bauer and Einfalt [112] found low reporting quality of these methods.

Brard et al [251] found evidence of poor reporting of Bayesian methods. To address this, when a Bayesian AD is used, authors should detail the model used for analysis to estimate the posterior probability distribution, the prior distribution used and rationale for its choice; whether the prior was updated in light of interim data and how; and clarify the stages when the prior information was used (interim or/and final analysis). If an informative prior was used, the source of data to inform this prior should be disclosed where applicable. Of note, part of the Bayesian community argue that it is not principled to control frequentist operating characteristics in Bayesian ADs [252], although these can be computed and presented [22,160,253].

Typically, ADs require quickly observed adaptation outcomes relative to the expected length of the trial. In some ADs, randomised participants who have received the treatment may not have their outcome data available at the interim analysis (referred to as overrunning participants) for various reasons [254]. These delayed responses may pose ethical dilemmas depending on the adaptive decisions taken, present logistical challenges, or diminish the efficiency of the AD depending on their prevalence and the objective of the adaptations [209]. It is therefore useful for readers to understand how overrunning participants were dealt with at interim analyses especially after a terminal adaptation decision (e.g., when a trial or treatment arms were stopped early for efficacy or futility). If outcome data of overrunning participants were collected, a description should be given of how these data were analysed and combined with interim results after the last interim decision was made. Some formal statistical methods to deal with accrued data from overrunning participants have been proposed [255].

Box 18. Exemplars on reporting item 12 elements

Example 1. GSD; statistical method for estimating treatment effects

"Stagewise ordering was used to compute the unbiased median estimate and confidence limits for the prognosis-group-adjusted hazard rates. [256]" [257]

Example 2. Inferentially seamless (4-arm 2-stage) AD with dose selection; statistical methods for controlling operating characteristics

" … the power of the study ranged from 71% to >91% to detect a treatment difference at a one-sided α of 0.025 when the underlying response rate of ≥1 of the crofelemer dose groups exceeded placebo by 20%. The clinical response of 20% was based on an estimated response rate of 55% in crofelemer and 35% in placebo during the 4-week placebo-controlled assessment period … For the primary endpoint, the test for comparing the placebo and treatment arms reflected the fact that data were gathered in an adaptive fashion and controlled for the possibility of an increased Type I error rate. Using the methods of Posch and Bauer, [71] as agreed upon during the special protocol assessment process, a p-value was obtained for comparison of each dose to the placebo arm from the stage I data, and an additional p-value was obtained for comparison of the optimal dose to the placebo arm from the independent data gathered in stage II. For the final primary analysis, the p-values from the first and second stages were combined by the inverse normal weighting combination function, and a closed testing procedure was implemented to test the null hypothesis using the methods of Posch and Bauer, [71] based on the original work of Bauer and Kieser [84]. This closed test controlled the experiment-wise error rate for this 2-stage adaptive design at a one-sided α of 0.025." [258] Extracted from appendix material.

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, [71,259,260] and the Hochberg-adjusted 1-sided P-value on stage 1 data. This testing procedure strongly controls the overall type I error rate at α level (see "Simulations run to assess the type I error rate under several null hypothesis scenarios"). Multiplicity-adjusted flexible repeated 95% 2-sided CIs [225] 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." [189] See supplementary material of the paper for details.

Example 4. Population-enrichment AD with SSR; criteria for claiming evidence of benefit; methods for controlling familywise type I error; combination test weights

Mehta et al. [104] published a methodological paper detailing a family of three hypotheses being tested; use of closure testing principle [260] to control the overall type I error; how evidence is claimed; and analytical derivations of the Simes adjusted p-values [261]. This include the use of combination test approach using pre-defined weights based on accrued information fraction for the 'full population' (cutaneous and non-cutaneous patients) and 'subpopulation' (cutaneous patients). Analytical derivations were presented for the two cases assuming enrichment occurs at interim analysis and no enrichment after interim analysis. Details are reported in a supplementary file accessible via the journal website.

Example 5. Inferentially seamless (7-arm 2-stage) AD with dose selection; use of traditional naïve estimates

"Unless otherwise stated, efficacy data are given as least squares means with standard error (SE) or 95% confidence interval (CI)." [83]

Example 6. Inferentially seamless phase 2/3 (5-arm 2-stage) AD with dose selection; dealing with overrunning participants

"Patients already assigned to an unselected regimen of propranolol by the time that the conclusions of the interim analysis are available, will continue the treatment according to the protocol but efficacy data for these patients will not be included in the primary analysis of primary endpoint." [103] Extracted from the supplementary material.

Results

CONSORT 2010 item 13a: For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome

ACE item 13a (modification): 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

Comment – Authors are referred to the CONSORT 2010 statement [3] for detailed discussion. Here, we only address additional requirements for ADs.

Explanation – The CONSORT 2010 statement [3,4] discusses why it is essential to describe participant flow adequately from screening to analysis. This applies to both interim and final analyses depending on the stage of reporting. The number of participants for each group with adaptation outcome data (that contributed to the interim analyses) should also be reported if different from the number of participants with primary outcome data. Furthermore, authors should report the number of randomised participants, for each group, that did not contribute to each interim analysis because of lack of mature outcome data at that interim look. For example, overrunning participants that were still being followed up when a terminal adaptation decision was made (e.g., dropping of treatment groups or early trial termination). The presentation of participant flow should align with the key hypotheses (e.g., subpopulation(s) and full study population) and treatment comparisons depending on the stage of results being reported.

Box 19. Exemplars on reporting item 13 (participant flowcharts)

Example 1. Inferentially seamless phase 2/3 AD

Additional file 2 is an illustrative structure that could be used to show the flow of participants when reporting the final results from a trial such as ADVENT [258].

Example 2. Population enrichment AD

Additional file 3 illustrates participant flowcharts that could be used for a population-enrichment adaptive trial such as the TAPPAS [104] which had key hypotheses relating to the cutaneous subpopulation and full population (cutaneous and non-cutaneous). The presentation depends on the trial adaptations made.

Example 3. Bayesian biomarker-targeted AD with RAR

Additional file 4 is an adapted flow diagram from BATTLE trial [262] showing the number of participants that contributed to the analysis by biomarker group during fixed randomisation (burn-in period) followed by RAR.

Section 14. Results (Recruitment)

CONSORT 2010 item 14a: Dates defining the periods of recruitment and follow-up

ACE item 14a (modification): Dates defining the periods of recruitment and follow-up, for each group

Comment – Authors should refer to the CONSORT 2010 statement [3,4] for the discussion.

Explanation – Consumers of research findings should be able to put trial results, study interventions, and comparators into context. Some ADs, such as those that evaluate multiple treatments allowing dropping of futile ones, selection of promising treatments, or addition of new treatments to an ongoing trial [19,263–265], incorporate pre-planned adaptations to drop or add new arms treatment arms during the course of the trial. As a result, dates of recruitment and follow-up may differ across treatment groups. In addition, the comparator arm may also change with time and concurrent or non-concurrent controls may be used. There are statistical implications that include how analysis populations for particular treatment comparisons are defined at different stages. For each treatment group, authors should clearly state the exact dates defining recruitment and follow-up periods. It should be stated if all treatment groups were recruited and followed-up during the same period.

Box 20. Exemplar on reporting item 14a

Example 1. MAMS platform AD

Figure 2 illustrates a graphical reporting of recruitment and follow-up periods for each treatment group including new arms that were added during the STAMPEDE trial. Corresponding comparator groups (controls) for treatment comparisons are indicated.

Fig 2 legend: Source – credit to Gilson et al. [266]. Reused in accordance with the terms of Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/). No changes to the original figure were made.

Example 2. Phase 2 Bayesian biomarker-targeted AD with RAR

"A total of 341 patients were enrolled in the BATTLE study between November 30, 2006, and October 28, 2009, with equally random assignments for the first 97 patients and adaptive randomization for the remaining 158." [262]

CONSORT 2010/ACE item 14b (clarification): Why the trial ended or was stopped

Comment – This item should be applied without reference to the CONSORT 2010 statement [3,4].

Explanation – Some clinical trials are stopped earlier than planned for reasons that will have implications for interpretation and generalisability of results. For example, poor recruitment is a common challenge [267]. This may limit the inference drawn or complicate interpretation of results based on insufficient or truncated trial data. Thus, the reporting of reasons for stopping a trial early including circumstances leading to that decision could help readers to interpret results with relevant caveats.

The CONSORT 2010 statement [3,4], however, did not distinguish early stopping of a trial due to a pre-planned adaptation from an unplanned change. To address this and for consistency, we have now reserved this item for reporting of reasons why the trial or certain treatment arm(s) were stopped outside the scope of pre-planned adaptations, including those involved in deliberations leading to this decision (e.g., sponsor, funder, or trial monitoring committee). We also introduced *item 14c* to capture aspects of adaptation decisions made in light of the accumulating data, such as stopping the trial or treatment arm because the decision-making criterion to do so has been met.

Box 21. Exemplar on reporting item 14b

Example 1. 2-stage AD with options for futility and efficacy early stopping and increase in sample size; unplanned trial termination
"The planned interim analysis of the study was done in November 2005 after 125 patients have been (sic) recruited… According to the adaptive design of the study, we therefore calculated another 163 patients per treatment group to be required to answer the primary question. Upon the slow accrual up to that timepoint, the study coordinators decided to close the trial at the end of 2005. Further analysis was regarded to be exploratory." [268]
Example 2. Sequential-step AD; unplanned trial termination
"… the third interim analysis indicated unexpectedly low initial cure rates in both arms; 84% in the multiple dose and 73% in the single-dose arm. The stopping rule was not met …, but based on the observed poor efficacy overall, and following discussions with the Data Safety and Monitoring Board (DSMB) and investigators, the sponsor terminated the trial." [269]

ACE item 14c (new): Specify what trial adaptation decisions were made in light of the pre-planned decision-making criteria and observed accrued data

Explanation – ADs depend on adherence to pre-planned decision rules to inform adaptations. Thus, it is vital for research consumers to be able to assess whether the adaptation rules were adhered to as pre-specified in the decision-making criteria given the observed accrued data at the interim analyses. Failure to adhere to pre-planned decision rules may undermine the integrity of the results and validity of the design by affecting the operating characteristics (see *item 7b* details on binding and non-binding decision rules).

Unforeseeable events can occur that may influence deviations from some pre-planned adaptation decisions rules (e.g., the overruling or ignoring certain rules). It is therefore essential to adequately describe which pre-planned adaptations were enforced, which were pre-planned but were not enforced or overruled even though the interim analysis decision rules indicated an adaptation should be made, and which unplanned changes were made other than unplanned early stopping of the trial or treatment arm(s) covered by *item 14b*. Pre-planned adaptations that were not implemented are difficult to assess because the interim decisions made versus the pre-planned intended decisions are often poorly reported, and reasons are rarely given [120]. The rationale for ignoring or overruling pre-planned adaptation decisions, or making unplanned decisions that affect the adaptations should be clearly stated and also who recommended or made such decisions (e.g., the data monitoring committee or adaptation committee). This enables assessment of potential bias in the adaptation decision-making process, which is crucial for the credibility of the trial.

Authors should indicate the point at which the adaptation decisions were made (i.e., stage of results) and any additional design changes that were made as a consequence of adaptation decisions (e.g., change in allocation ratio).

Box 22. Exemplar on reporting item 14c elements

Example 1. Bayesian adaptive-enrichment AD with futility and superiority early stopping; stage of results
"Enrolment in the trial was stopped at 31 months, because the results of an interim analysis met the pre-specified criterion for trial discontinuation, which was a predictive probability of superiority of thrombectomy of at least 95% for the first primary endpoint (the mean score for disability on the utility-weighted modified Rankin scale at 90 days). This was the first pre-specified interim analysis that permitted stopping for this reason, and it was based on the enrolment of 200 patients. Because enrichment thresholds had not been crossed, the analysis included the full population of patients enrolled in the trial, regardless of infarct volume." [105]
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." [145]

Section 15. Results (Baseline data)

CONSORT 2010 item 15: A table showing baseline demographic and clinical characteristics for each group

ACE Item 15a «15 (clarification, renumbered): A table showing baseline demographic and clinical characteristics for each group

Comments – We renumbered the item to accommodate the new *item 15b*. This item should be applied with reference to the CONSORT 2010 statement [3,4], with additional requirements for specific ADs.

Explanation – The presentation of treatment group summaries of key characteristics and demographics of randomised participants who contributed to results influence interpretation, and helps readers and medical practitioners to make judgements about which patients the results are applicable to. For some ADs,

such as population (biomarker or patient) enrichment [91,151] when the study population is considered heterogeneous, a trial could be designed to evaluate if study treatments are effective in specific pre-specified sub-populations or a wider study population (full population). A pre-planned adaptation strategy may involve testing the effect of treatments in both pre-specified sub-populations of interest and the wider population in order to target patients likely to benefit the most. For such ADs, it is essential to provide summaries of characteristics of those who were randomised and who contributed to the results being reported (both interim or final), by treatment group for each subpopulation of interest and the full population consistent with hypotheses tested. These summaries should be reported without hypothesis testing of baseline differences in participants characteristics because it is illogical in randomised trials [270–273]. The CONSORT 2010 statement [3,4] presents an example of how to summarise baseline characteristics.

In the presence of marked differences in the number of randomised participants and those included in the interim or final analyses, authors are encouraged to report baseline summaries by treatment group for these two populations. Readers will then be able to assess representativeness of the interim or final analysis population relative to those randomised and also the target population.

Box 24. Exemplar on reporting item 15a
Example. Population-enrichment AD
See Additional file 5 for a dummy baseline table for the TAPPAS trial [104].

ACE item 15b (new): Summary of data to enable the assessment of similarity in the trial population between interim stages

Comment – This item is applicable for ADs conducted in distinct stages for which the trial has progressed beyond the first stage.

Explanation – Changes in trial conduct and other factors may introduce heterogeneity in the characteristics or standard management of patients before and after trial adaptations. Consequently, results may be inconsistent or heterogeneous between stages (interim parts) of the trial [209]. For ADs, access to interim results or mere guesses based on interim decisions taken may influence behaviour of those directly involved in the conduct of the trial and thus introduce operational bias [22]. Some trial adaptations may introduce intended changes to inclusion or exclusion criteria (e.g., population enrichment [96,151]). Unintended changes to characteristics of patients over time may occur (population drift) [274]. A concern is whether this could lead to a trial with a different study population that does not address the primary research objectives [275]. This jeopardises validity, interpretability, and credibility of trial results. It may be difficult to determine whether differences in characteristics between stages occured naturally due to chance, were an unintended consequence of pre-planned trial adaptations, represent operational bias introduced by knowledge or communication of interim results, or are for other reasons [276]. However, details related to *item 11c* may help readers make informed judgements on whether any observed marked differences in characteristics between stages are potentially due to systematic bias or just chance. Therefore, it is essential to provide key summary data of participants included in the analysis (as discussed in *item 15a*) for each interim stage of the trial and overall. Authors are also encouraged to give summaries by stage and treatment group. This will help readers assess similarity in the trial population between stages and whether it is consistent across treatment groups.

Box 25. Exemplar on reporting item 15b elements

Example. Overall baseline characteristics by stage; inferentially seamless phase 2/3 AD

Table 5: Characteristics of randomised participants (N=1202) in stage 1 and 2.

Characteristic	Stage 1 (n=230)	Stage 2 (n=972)
Age (years), mean (SD)	53.4 (10.3)	54.3 (9.7)
Gender (female), n (%)	139 (60.4)	504 (51.9)
Race (white), n (%)	103 (44.8)	509 (52.4)
BMI (kg/m^2), mean (SD)	31.9 (4.5)	31.1 (4.3)
Body weight (kg), mean (SD)	87.3 (18.0)	86.2 (17.1)
Duration of diabetes (years), mean (SD)	7.5 (5.5)	7.0 (5.1)
Seated systolic BP (mm Hg), mean (SD)	128.0 (14.4)	127.7 (13.1)
Seated diastolic BP (mm Hg), mean (SD)	77.9 (7.9)	77.6 (8.6)
Seated heart rate (bpm), mean (SD)	74.5 (9.6)	75.2 (10.0)

Adapted from Geiger et al. [172]; 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.

Section 16. Results (Number analysed)

CONSORT 2010/ACE item 16 (clarification): For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups

Comment – The item should be used in reference to the CONSORT 2010 statement [3,4] for original details and examples. Here, we give additional clarification for some specific requirements of certain ADs such as population enrichment [91,151].

Explanation – We clarify that the number of participants by treatment group should be reported for each analysis at both the interim analyses and final analysis whenever a comparative assessment is performed (e.g., for efficacy, effectiveness, or safety). Most importantly, the presentation should reflect the key hypotheses considered to address the research questions. For example, population (patient or biomarker) enrichment AD can be reported by treatment group for each pre-specified subpopulation and full population depending on key hypotheses tested.

Section 17. Results (Outcomes and estimation)

CONSORT 2010/ACE item 17a (clarification): For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)

Comment – We expanded the explanatory text to address some specific requirements of certain ADs such as population enrichment [151]. Therefore, the item should be used in reference to the CONSORT 2010 [3] for original details and examples.

Explanation – In randomised trials, we analyse participant outcome data collected after study treatments are administered to address research questions about beneficial and/or harmful effects of these treatments. In principle, reported results should be in line with the pre-specified estimand(s) and compatible with the research questions or objectives [214,215]. The CONSORT 2010 statement [3,4] addresses what authors should report depending on the outcome measures. These include group summary measures of effect, for both interim and final analyses, including the number of participants contributing to the analysis, appropriate measures of the treatment effects (e.g., between group effects for a parallel group randomised trials) and associated uncertainty (e,g credible or confidence intervals). Importantly, the presentation is influenced by how the key hypotheses are configured to address the research questions. For some ADs, such as population/patient enrichment, key hypotheses often relate to whether the study treatments are effective in the whole target population of interest or in specific subpopulations of the target population classified by certain characteristics. In such ADs, reporting of results as detailed in the CONSORT 2010 should mirror hypotheses of interest. That is, we expect the outcome results to be presented for the subpopulations and full target populations considered by treatment group. This is to help readers interpret results on whether the study treatments are beneficial to the target population as a whole or only to specific pre-specified subpopulations.

ACE item 17c (new): Report interim results used to inform interim decision-making

Explanation – Adherence to pre-planned adaptations and decision rules including timing and frequency is essential in AD randomised trials. This can only be assessed when the pre-planned adaptations (*item 3b*), adaptive decision rules (*item 7b*), and results that are used to guide the trial adaptations are transparently and adequately reported.

Marked differences in treatment effects between stages may arise (e.g., discussed in *item 15b*) making overall interpretation of their results difficult [96,115,274,276–279]. The presence of heterogeneity questions the rationale for combining results from independent stages to produce overall evidence, as is also the case for combining individual studies in a meta-analysis [96,280]. Although this problem is not unique to AD randomised trials, consequences of trial adaptation may worsen the problem [276]. Authors should at least report the relevant interim or stage results that were used to make each adaptation, consistent with *items 3b* and *7b*; for example, interim treatment effects with uncertainty, interim conditional power or variability used for SSR, and trend in the probabilities of allocating participants to a particular treatment group as the trial progresses. Authors should report interim results of treatment groups or subpopulations that have been dropped due to lack of benefit or poor safety. This reduces the reporting bias caused by selective disclosure of treatments only showing beneficial and/or less harmful effects.

Box 26. Exemplars on reporting item 17 elements

Example 1. Bayesian RAR, change in randomisation probabilities across arms throughout the trial; randomisation updates were made after every patient

Fig 3 legend: Source – credit to Pallmann et al [22]. Reused in accordance with the terms of Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/). No changes to the original figure were made.

Giles et al. [73] present a table of changes in allocation probabilities used to create Figure 3 by treatment group including allocated treatment and primary outcome response for each participant.

Example 2. Inferentially seamless phase 2/3 AD; stage 1 treatment selection results

Barnes et al. [145] clearly presented the results that led to the interim selection of the two indacaterol drug doses to progress to stage 2 of the study; 150µg (the lowest dose that exceeded both pre-specified treatment selection criteria) and 300µg (the next highest dose that met the same criteria). The interim difference in treatment effect compared to placebo with uncertainty per group for the two adaptation outcomes are displayed in Figures 1 and 2 of the paper.

Example 3. 2-stage GSD; stage 1 dose selection results

"At the interim analysis planned after at least 1969 patients had been randomized and reached day 7 follow-up in each group,[189] the otamixaban dose for stage 2 of the trial was selected as described in eFigure 1 in the Supplement. At that time, the rates of the primary efficacy outcome in the higher-dose otamixaban group was xx (4.7%) (the one selected to go forward) and was xx (5.6%) in the UFH-pluseptifibatide group (adjusted RR, 0.848; 95% CI, 0.662-1.087) but the lower-dose group fulfilled the pre-specified criteria for futility with a RR of more than 1 (primary efficacy outcome, xx (6.3%); RR, 1.130; 95% CI, 0.906-1.408) and was discontinued." [148] xx are the corresponding number of participants with primary response that should have been stated.

Example 4. Adapted from Khalil et al [147]; sequential-step AD

Table 6. Interim results

Parasite clearance at day 30 (initial cure)	Treatment group	Parasite clearance rate, n/N (%)	Differences in parasite clearance rates (95% CI)	P-value
Interim analysis 1	Single dose, 7.5mg/kg	10/20 (50.0%)	Reference	
	Multiple dose, 7x3mg/kg	16/18 (88.9%)	38.9% (12.6 to 65.2)	0.015 [a]
Interim analysis 2 [b]	Single dose, 10mg/kg	16/20 (80.0%)	Reference	
	Multiple dose, 7x3mg/kg	19/25 (76.0%)	-4.0% (-28.2 to 20.2)	0.748 [c]
Interim analysis 3[d]	Single dose, 10mg/kg	29/40 (72.5%)	Reference	
	Multiple dose, 7x3mg/kg	37/44 (84.1%)	11.6% (-6.0 to 29.1)	0.196 [e]

N, total number of patients per group (denominator); n, patients with recorded parasitic clearance per groups (events); CI, confidence interval; [a] p-value from Fisher's exact test, adaptation rule met to escalate dose so dosage increased to 10 mg/kg and continue recruitment; [b] adaptation rule to escalate dose not met so recruitment was continued with the same dosage (10mg/kg in single-dose arm; [c, e] p-values from a Chi-square test; [e] adaptation rule to escalate dose not met but concerns arose regarding low cure in each arm and recruitment was terminated; [d] includes patients in interim analysis 2; patients in interim analysis 1 did not contribute to any subsequent interim analysis.

Discussion

Discussion (Limitations)

CONSORT 2010/ACE item 20 (clarification): Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses

Comments – No change in wording is made to this item so it should be applied with reference to the CONSORT 2010 statement [3,4] for original details and examples. Here, we only address additional considerations for ADs.

Explanation – We expect authors to discuss the arguments for and against the implemented study design and its findings. Several journals have guidelines for structuring the discussion to prompt authors to discuss key limitations with possible explanations. The CONSORT 2010 statement [3,4] addresses general aspects relating to potential sources of bias, imprecision, multiplicity of analyses and implications of unplanned changes to methods or design. For AD randomised trials, further discussion should include the implications of:

- any deviations from the pre-planned adaptations (e.g., decision rules that were not enforced or overruled and changes in timing or frequency of interim analyses);
- interim analyses (e.g., updating randomisation with inadequate burn-in period);
- protocol amendments on the trial adaptations and results;
- potential sources of bias introduced by interim analyses or decision-making; potential bias and imprecision of the treatment effects if naïve estimation methods were used;
- potential heterogeneity in patient characteristics and treatment effects between stages;
- maturing of efficacy and safety data to robustly inform trial adaptations at interim analyses and;
- using adaptation outcome(s) different from the primary outcome(s).

Additionally, it is encouraged to discuss the observed efficiencies of pre-planned adaptations in addressing the research questions and lessons learned about using the AD, both negative and positive. This is optional as it does not directly influence the interpretation of the results but enhances much-needed knowledge transfer of innovative trial designs. Therefore, authors have been encouraged to consider separate methodology publications in addition to trial results [59,281].

Box 27. Exemplars on reporting item 20

Example 1. Use of surrogate outcome to inform adaptation
“We chose change in the SOFA scores as a surrogate outcome based on strong correlations between this measure and 28-day mortality (33). Whether schange in the SOFA scores and the timing of reassessment (48 hours in this case) represents the “right” surrogate endpoint for nonpivotal sepsis trials remains unclear and is an area for future consideration, although the use of change in the SOFA score as a surrogate outcome is supported by a recent meta-analysis (34).” [146]
Example 2. Duration of assessments to inform dose selection
“The use of the adaptive seamless design is not without potential risk. The initial dose-finding period needs to be long enough for a thorough evaluation of effects. Two weeks was considered a fully adequate period in which to attain pharmacodynamic steady state …” [145]
Example 3. Early stopping outside the scope of the pre-planned adaptation and possible explanation
“The aim was to determine the minimum efficacious dose and safety of treatments in HIV-uninfected patients. However, the study had to be prematurely terminated due to unacceptably low efficacy in both the single and multiple dose treatment arms, with a cure rate of only 85% in the multiple-dose arm. Adverse effects of treatment in this study were in line with the current drug label. The overall low efficacy was unexpected, as total doses of 10 mg/kg and above resulted in DC rates of at least 90% in a trial in Kenya (13). The trial was not powered for data analysis by geographical location (centre) and the results may have been due to chance, but both the 10 mg/kg single dose and 21 mg/kg multiple dose regimens appeared to work very well in the small number of patients treated in Arba Minch Hospital (southern Ethiopia). We have little explanation for the overall poor response seen in this study or for the observed geographical variations. Previously, similar geographical variation in treatment response in these three sites was seen for daily doses of 11 mg/kg body weight paromomycin base over 21 days (7), a regimen which had also proven efficacious in India (18). Methodological bias is unlikely in this randomized trial, but differences in base line patient characteristics between the three trial sites could have possibly introduced bias, leading to variation in treatment response …” [147]
Example 4. Limitations of biomarkers and RAR
“Our study has some important limitations. First, and probably most important, our biomarker groups were less predictive than were individual biomarkers, which diluted the impact of strong predictors in determining treatment probabilities. For example, EGFR mutations were far more predictive than was the overall EGFR marker group. The unfortunate decision to group the EGFR markers also impacted the other marker groups and their interactions with other treatments, resulting in a suboptimal overall disease control rate as described. Second, several of the pre-specified markers (e.g., RXR) had little, if any, predictive value in optimizing treatment selections. This limitation will be addressed in future studies by not grouping or prespecifying biomarkers prior to initiating these biopsy-mandated trials. In addition, adaptive randomization, which assigns more patients to the more effective treatments within each biomarker group, only works well with a large differential efficacy among the treatments (as evident in the KRAS/BRAF group), but its role is limited without such a difference (e.g., in the other marker groups). Allowing prior use of erlotinib was another limitation and biased treatment assignments; in fact, the percentage of patients previously treated with erlotinib steadily increased during trial enrollment. Overall, 45% of our patients were excluded from the 2 erlotinib-containing arms because of prior EGFR TKI treatment. As erlotinib is a standard of care therapy in NSCLC second-line, maintenance, and front-line settings, the number of patients receiving this targeted agent will likely continue to increase.” [262]

Section 21. Discussion (Generalisability)

CONSORT 2010/ACE item 21 (clarification): Generalisability (external validity, applicability) of the trial findings

Comments – We have not changed the wording of this item so it should be considered in conjunction with the CONSORT 2010 statement [3,4]. However, there are additional considerations that may influence the generalisability of results from AD randomised trials.

Explanation – Regardless of the trial design, authors should discuss how the results are generalisable to other settings or situations (external validity) and how the design and conduct of the trial minimised or mitigated potential sources of bias (internal validity) [3]. For ADs, there are many factors that may undermine both internal (see *item 20* clarifications) and external validity. Trial adaptations are planned with a clear rationale to achieve research goals or objectives. Thus, the applicability of the results may be intentionally relevant to the target population enrolled or pre-specified subpopulation(s) with certain characteristics (subsets of the target population). Specifically, the implemented adaptations and other factors may cause unintended population drift or inconsistencies in the conduct of the trial. Authors should discuss the population to whom the results are applicable including any threats to internal and external validity which are trial dependent based on the implemented adaptations.

Box 28. Exemplar on reporting item 21 elements

Example 1. Bayesian population-enrichment AD with RAR; to whom the results are applicable (full population)
"The DAWN trial showed that, among patients with stroke due to occlusion of the intracranial internal carotid artery or proximal middle cerebral artery who had last been known to be well 6 to 24 hours earlier and who had a mismatch between the severity of the clinical deficit and the infarct volume, outcomes for disability and functional independence at 90 days were better with thrombectomy plus standard medical care than with standard medical care alone." [105]
Example 2. Phase 2 Bayesian biomarker-targeted AD with RAR; to whom the results are applicable (biomarker specific)
"Sorafenib was active against tumors with mutated or wild-type KRAS, but had a worse disease control rate (compared with other study agents) in patients with EGFR mutations. As expected (5–7, 15–17), erlotinib was beneficial in patients with mutated-EGFR tumors. Erlotinib plus bexarotene improved disease control in patients with a higher expression of Cyclin D1, suggesting a potential role for bexarotene in lung cancer treatment (11); similar to sorafenib, the combination also improved disease control in the KRAS-mutant patient population. Future randomized, controlled studies are needed to further confirm the predictive value of these biomarkers." [262] Liu and Lee [89] published details of the design and conduct of this trial.

Section 24. Other information (Statistical analysis plan and other relevant trial documents)

ACE item 24b (new): Where the full statistical analysis plan and other relevant trial documents can be accessed

Explanation – Pre-specifying details of statistical methods and their execution including documentation of amendments, and when they occurred is good scientific practice that enhances trial credibility and reproducibility of methods, results and inference. The Statistical Analysis Plan (SAP) is the principal technical document that details the statistical methods for the design of the study; analysis of the outcomes; aspects that influence the analysis approaches; and presentation of results consistent with the research questions/objectives and estimands [214,215] in line with the trial protocol (now *item 24a*). General guidance on statistical principles for clinical trials to consider with the aim to standardise research practice exist [282–284]. AD trials tend to bring additional statistical complexities and considerations during the design and analyses depending on the trial adaptations considered. Access to the full SAP with amendments addressing interim and final analyses is essential. This can be achieved through the use of several platforms such as online supplementary material, online repositories, or referencing published material. This enables readers to access additional information relating to the statistical methods that may not be feasible to include in the main report.

Critical details of the trial adaptations (e.g., the decision-making criteria or adaptation algorithm and rules) may be intentionally withheld from publicly accessible documents (e.g., protocol) while the trial is ongoing [46,211]. These details may be documented in a formal document with restricted access and disclosed only when the trial is completed in order to minimise operational bias (*item 11c*). For this situation, authors should provide access to such details withheld with any amendments made for transparency and an audit trail of pre-planned AD aspects.

For some AD randomised trials, methods to derive statistical properties analytically may not be available. Thus, it becomes necessary to perform simulations under a wide range of plausible scenarios to investigate the operating characteristics of the design (*item 7a*), impact on estimation bias (*item 12b*), and appropriateness and consequences of decision-making criteria and rules [160,285]. In such cases, we encourage authors to reference accessible material used for this purpose (e.g., simulation protocol and report, or published related material). Furthermore, it is a good scientific practice to reference software, programs or code used for this task to facilitate reproducible research.

The operating characteristics of ADs heavily depend on following the pre-planned adaptations and adaptive decision-making criteria and rules. ADs often come with additional responsibilities for the traditional monitoring committees or require a specialised monitoring committee to provide independent oversight of the trial adaptations (e.g., adaptive decision-making or adaptation committee). Thus, it is essential to be transparent about the adaptation decision-making process, roles and responsibilities of the delegated DMC(s), recommendations made by the committee and whether recommendations were adhered to. Authors are encouraged to provide supporting evidence (e.g., DMC charter).

Box 29. Exemplars on reporting item 24b
Example 1. Interim and final SAPs; IDMC roles and responsibilities; supplementary material
Léauté-Labrèze et al. [103] provide several versions of the SAP for a two-stage inferentially seamless phase 2/3 AD as supplementary material. The remit and responsibilities of the IDMC including involvement in the adaptation decision-making process are detailed. The last version (3.5) of the SAP with amendments and details of interim and final analysis is found on page 759 to 830 of the protocol supplementary material. Simulation results are summarised on pages 831 to 836.
Example 2. Simulation report; supplementary material
Steg et al. [189] provide a simulation report evaluating the operating characteristics of a 3-arm 2-stage group sequential AD with dose selection under a number of scenarios in an appendix. The authors also explored the bias in methods used to estimate the treatment effects and confidence intervals and used the simulation results to inform their choice of methods.
Example 3. Set-up of simulation studies and simulation results; published methodology work
Gu et al. [216] describe how simulation studies were performed and presented simulation results for evaluating operating characteristics of a two-stage Bayesian biomarker-based AD.
Example 4. Simulation report; published methodology work
Skrivanek et al. [160] published extensive simulation work quantifying operating characteristics of a Bayesian inferentially seamless phase 2/3 AD with RAR.
Example 5. Simulation report; published methodology work
Heritier et al. [69] published extensive simulation work for an inferentially seamless phase 2/3 design using frequentist methods.

Conclusions

There is a multidisciplinary desire to improve efficiency in the conduct of randomised trials. ADs allow pre-planned adaptations that offer opportunities to address research questions in randomised trials more efficiently compared to fixed designs. However, ADs can make the design, conduct and analysis of trials more complex. Potential biases can be introduced during the trial in several ways. Consequently, there are additional demands for transparency and reporting to enhance the credibility and interpretability of results from adaptive trials.

This CONSORT extension provides minimum essential reporting requirements that are applicable to pre-planned adaptations in randomised AD trials, designed and analysed using Frequentist or Bayesian statistical methods. We have also given many exemplars of different types of ADs to help authors when using this extension. Our consensus process involved stakeholders from the public and private sectors [13,136]. We hope this extension will facilitate better reporting of randomised ADs and indirectly improve their design and conduct, as well as much-needed knowledge transfer.

Abbreviations

ACE, Adaptive designs CONSORT Extension

AD, adaptive design

CONSORT, Consolidated Standards Of Reporting Trials

E&E, explanation and elaboration

EQUATOR, Enhancing the QUAlity and Transparency Of health Research

(I)DMC, (independent) data monitoring committee

GSD, group sequential design

MAMS, multi-arm multi-stage design

MeSH, Medical Subject Heading

RAR, response-adaptive randomisation

SAP, Statistical Analysis Plan

SSR, sample size re-estimation/re-assessment /re-calculation

Declarations

Ethics approval and consent to participate

The project ethics approval was granted by the Research Ethics Committee of the School of Health and Related Research (ScHARR) at the University of Sheffield (ref: 012041). All Delphi participants during the development of the ACE guideline provided consent online during survey registration.

Consent for publication

Not applicable

Availability of data and material

Not applicable. We have published the development process of this guideline including anonymised participant data from Delphi surveys which is publicly accessible [13].

Competing interests

MKW is an employee of Janseen Research and Development, Inc. All other authors declare no competing interests.

Funding

This paper summarises independent research jointly funded by the National Institute for Health Research (NIHR) Clinical Trials Unit Support Funding scheme and the Medical Research Council (MRC) Network of Hubs for Trials Methodology Research (HTMR) (MR/ L004933/-2/N90).

The funders had no role in the design of the study and collection, analysis, and interpretation of data and in writing or approving the manuscript.

MD, JW, TJ, APM, ST, SAJ, FK, CJW, DH, and JN were co-applicants who sourced funding from the NIHR CTU Support Funding scheme. MD, JW and TJ sourced additional funding from the MRC HTMR.

TJ's contribution was in part funded by his Senior Research Fellowship (NIHR-SRF-2015-08-001) funded by NIHR. NHS Lothian via the Edinburgh Clinical Trials Unit supported CW in this work. The University of Sheffield via the Sheffield Clinical Trials Unit and the NIHR CTU Support Funding scheme supported MD's contribution. An MRC HTMR grant (MR/L004933/1-R/N/P/B1) partly funded PP's contribution.

The views expressed are those of the authors and not necessarily those of the National Health Service (NHS), the NIHR, the MRC, or the Department of Health and Social Care.

This work reflects the views of the authors and should not be construed to represent the US Food and Drug Administration's views or policies

Views expressed in this paper are the author's professional opinions and do not necessarily represent the official positions of the Pharmaceutical and Medical Devices Agency, Japan

Authors' contributions

The idea originated from MD's NIHR Doctoral Research Fellowship (DRF-2012-05-182) under the supervision of SAJ, ST, and JPN. The idea was presented, discussed, and contextualised at the 2016 annual workshop of the MRC Network of HTMR Adaptive Designs Working Group (ADWG), which was attended by six members of the Steering Committee (MD, TJ, PP, JW, APM, and CJW). MD, JW, TJ, APM, ST, SAJ, FK, CJW, DH, and JPN conceptualised the study design and applied for funding. All authors contributed to the conduct of the study and interpretation of the results. MD, ST, PP, CJW, JW, TJ and SAJ led the write-up of the manuscript. All authors except DGA contributed to the write-up of the manuscript, reviewed, and approved this final version. DGA oversaw the development process on behalf of the CONSORT Group and also contributed to the initial draft manuscript until his passing after the approval of the final checklists. The Steering Committee is deeply saddened by the passing of DGA who did not have the opportunity to approve the final manuscript. We dedicate this work to him in memory of his immense contribution to the ACE project, medical statistics, good scientific research practice and reporting, and humanity.

Members of the ACE Consensus Group contributed to the consensus workshop, approved the final ACE checklist, and were offered the opportunity to review the final manuscript.

Acknowledgements

ACE Consensus Group

We are very grateful to the consensus meeting participants who provided recommendations on reporting items to retain in this guidance and for comments on the structure of this E&E document. All listed members approved the final ACE checklist and offered opportunity to review this E&E document. Sally Hopewell joined the group during the write up stage to oversee the process on behalf of the CONSORT Group following the sad passing of Doug G. Altman.

Munyaradzi Dimairo [1]; Toshimitsu Hamasaki [2]; Susan Todd [3]; Christopher J. Weir [4]; Adrian P. Mander [5]; James Wason [5, 6]; Franz Koenig [7]; Steven A. Julious [8]; Daniel Hind [1]; Jon Nicholl [1]; Douglas G. Altman [9]; William J. Meurer [10]; Christopher Cates [11]; Matthew Sydes [12]; Yannis Jemiai [13]; Deborah Ashby [14] (Chair); Christina Yap [15]; Frank Waldron-Lynch [16]; James Roger [17]; Joan Marsh [18]; Olivier Collignon [19]; David J. Lawrence [20]; Catey Bunce [21]; Tom Parke [22]; Gus Gazzard [23]; Elizabeth Coates [1]; Marc K Walton [25]; and Sally Hopewell [9].

[1] School of Health and Related Research, University of Sheffield, Sheffield, UK

[2] National Cerebral and Cardiovascular Center, Osaka, Japan

[3] Department of Mathematics and Statistics, University of Reading, Reading, UK

[4] Edinburgh Clinical Trials Unit, Centre for Population Health Sciences, Usher Institute of Population Health Sciences & Informatics, The University of Edinburgh, Edinburgh, UK

[5] MRC Biostatistics Unit, University of Cambridge, School of Clinical Medicine, Cambridge Institute of Public Health, Cambridge, UK

[6] Institute of Health and Society, Newcastle University, UK

[7] Medical University of Vienna, Center for Medical Statistics, Informatics, and Intelligent Systems, Vienna, Austria

[8] Medical Statistics Group, School of Health and Related Research, University of Sheffield, Sheffield, UK

[9] Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology & Musculoskeletal Sciences, University of Oxford, Oxford, UK

[10] Departments of Emergency Medicine and Neurology, University of Michigan, Ann Arbor, MI, USA

[11] Cochrane Airways, PHRI, SGUL, London, UK

[12] MRC Clinical Trials Unit, UCL, Institute of Clinical Trials & Methodology, London, UK

[13] Cytel, Cambridge, USA

[14] School of Public Health, Imperial College London, UK

[15] Cancer Research UK Clinical Trials Unit, Institute of Cancer and Genomic Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

[16] Novartis Institutes for Biomedical Research, Basel, Switzerland

[17] Institutional address not applicable

[18] The Lancet Psychiatry, London, UK

[19] Luxembourg Institute of Health, Strassen, Luxembourg

[20] Novartis Campus, Basel, Switzerland

[21] School of Population Health and Environmental Sciences, Faculty of Life Sciences and Medicine, King's College London, London, UK

[21] Berry Consultants, Merchant House, Abingdon, UK

[23] NIHR Biomedical Research Centre at Moorfields Eye Hospital and UCL Institute of Ophthalmology, London, UK

[25] Janssen Research and Development, Ashton, USA

ACE Steering Committee

Munyaradzi Dimairo, Elizabeth Coates, Philip Pallmann, Susan Todd, Steven A. Julious, Thomas Jaki, James Wason, Adrian P. Mander, Christopher J. Weir, Franz Koenig, Marc K. Walton, Katie Biggs, Jon P. Nicholl, Toshimitsu Hamasaki, Michael A. Proschan, John A. Scott, Yuki Ando, Daniel Hind, and Douglas G. Altman

External Expert Panel

We would like to thank William Meurer [a], Yannis Jemiai [b], Stephane Heritier [c], and Christina Yap [d] for their invaluable contributions in reviewing the checklist and working definitions of technical terms.

[a] University of Michigan, Departments of Emergency Medicine and Neurology, Ann Arbor, MI, USA

[b] Cytel, Cambridge, USA

[c] Monash University, Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Australia

[d] Cancer Research UK Clinical Trials Unit, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK

Administrative support

We thank Sarah Gonzalez for coordination and administrative support throughout the project, Sheffield Clinical Trials Research Unit for their support, particularly Mike Bradburn and Cindy Cooper, and Benjamin Allin and Anja Hollowell for the Delphi surveys technical and administrative support.

Delphi survey participants

We are grateful for the input of several key stakeholders during the Delphi surveys and many who gave informal advice.

Other support

Trish Groves contributed to the consensus workshop, reviewed and approved the finalised ACE checklist. Lingyun Liu provided dummy baseline information for the TAPPAS trial.

References

1. Yordanov Y, Dechartres A, Porcher R, Boutron I, Altman DG, Ravaud P. Avoidable waste of research related to inadequate methods in clinical trials [Internet]. BMJ. British Medical Journal Publishing Group; 2015 [cited 2018 Jun 12]. p. h809. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25804210
2. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of evidence. Lancet [Internet]. Elsevier; 2009 [cited 2018 Jun 12];374:786. Available from: https://www.sciencedirect.com/science/article/pii/S0140673609603299?via%3Dihub
3. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ [Internet]. 2010 [cited 2014 Jul 10];340:c869. Available from: http://www.bmj.com/content/340/bmj.c869

4. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med [Internet]. 2010 [cited 2014 Nov 24];152:726–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20335313
5. CONSORT Group. Extensions of the CONSORT Statement [Internet]. [cited 2019 Mar 14]. Available from: http://www.consort-statement.org/extensions
6. Ivers NM, Taljaard M, Dixon S, Bennett C, McRae A, Taleban J, et al. Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000-8. BMJ [Internet]. British Medical Journal Publishing Group; 2011 [cited 2018 Jan 28];343:d5886. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21948873
7. Moher D, Jones A, Lepage L. Use of the CONSORT Statement and Quality of Reports of Randomized Trials. JAMA [Internet]. American Medical Association; 2001 [cited 2014 Nov 24];285:1992. Available from: http://jama.jamanetwork.com/article.aspx?articleid=193739
8. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review [Internet]. Med. J. Aust. 2006 [cited 2018 Sep 23]. p. 263–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16948622
9. Blanco D, Biggane AM, Cobo E, Altman D, Bertizzolo L, Boutron I, et al. Are CONSORT checklists submitted by authors adequately reflecting what information is actually reported in published papers? [Internet]. Trials. 2018 [cited 2018 Sep 23]. p. 80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29378640
10. Jin Y, Sanger N, Shams I, Luo C, Shahid H, Li G, et al. Does the medical literature remain inadequately described despite having reporting guidelines for 21 years? - A systematic review of reviews: an update. J Multidiscip Healthc [Internet]. Dove Press; 2018 [cited 2018 Oct 16];11:495–510. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30310289
11. Janackovic K, Puljak L. Reporting quality of randomized controlled trial abstracts in the seven highest-ranking anesthesiology journals. Trials [Internet]. BioMed Central; 2018 [cited 2018 Oct 30];19:591. Available from: https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-018-2976-x
12. Goldacre B, Drysdale H, Dale A, Milosevic I, Slade E, Hartley P, et al. COMPare: a prospective cohort study correcting and monitoring 58 misreported trials in real time. Trials [Internet]. BioMed Central; 2019 [cited 2019 Feb 18];20:118. Available from: https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-019-3173-2
13. Dimairo M, Coates E, Pallmann P, Todd S, Julious SA, Jaki T, et al. Development process of a consensus-driven CONSORT extension for randomised trials using an adaptive design. BMC Med [Internet]. BioMed Central; 2018 [cited 2018 Nov 16];16:210. Available from: https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-018-1196-2
14. FDA. Adaptive Designs for Medical Device Clinical Studies: Draft Guidance for Industry and Food and Drug Administration Staff [Internet]. 2015. Available from: http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm446729.pdf
15. Chow S-C, Chang M. Adaptive design methods in clinical trials - a review. Orphanet J Rare Dis [Internet]. 2008 [cited 2014 Jul 10];3:11. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2422839&tool=pmcentrez&rendertype=abstract
16. Dragalin V. Adaptive Designs: Terminology and Classification. Drug Inf J. 2006;40:425–35.
17. Gallo P, Chuang-Stein C, Dragalin V, Gaydos B, Krams M, Pinheiro J. Adaptive designs in clinical drug development--an Executive Summary of the PhRMA Working Group. J Biopharm Stat [Internet]. 2006 [cited 2014 Nov 24];16:275-83; discussion 285-91, 293–8, 311–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16724485
18. Kairalla J a, Coffey CS, Thomann M a, Muller KE. Adaptive trial designs: a review of barriers and opportunities. Trials [Internet]. Trials; 2012 [cited 2014 Nov 17];13:145. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3519822&tool=pmcentrez&rendertype=abstract
19. Lewis RJ. The pragmatic clinical trial in a learning health care system. Clin Trials [Internet]. 2016;13:484–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27365011
20. Curtin F, Heritier S. The role of adaptive trial designs in drug development [Internet]. Expert Rev. Clin. Pharmacol. Taylor & Francis; 2017 [cited 2017 Apr 29]. p. 727–36. Available from: https://www.tandfonline.com/doi/full/10.1080/17512433.2017.1321985
21. Park JJ, Thorlund K, Mills EJ. Critical concepts in adaptive clinical trials. Clin Epidemiol [Internet]. Dove Press; 2018 [cited 2018 Mar 30];10:343–51. Available from: https://www.dovepress.com/critical-concepts-in-adaptive-clinical-trials-peer-reviewed-article-CLEP
22. Pallmann P, Bedding AW, Choodari-Oskooei B, Dimairo M, Flight L, Hampson L V., et al. Adaptive designs in clinical trials: why use them, and how to run and report them. BMC Med [Internet]. BioMed Central; 2018 [cited 2018 Mar 1];16:29. Available from: https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-018-1017-7
23. Jaki T, Wason JMS. Multi-arm multi-stage trials can improve the efficiency of finding effective treatments for stroke : a case study and review of the literature. BMC Cardiovasc Disord [Internet]. BioMed Central; 2018 [cited 2018 Dec 6];18:215. Available from: https://bmccardiovascdisord.biomedcentral.com/articles/10.1186/s12872-018-0956-4
24. Parmar MK, Sydes MR, Cafferty FH, Choodari-Oskooei B, Langley RE, Brown L, et al. Testing many treatments within a single protocol over 10 years at MRC Clinical Trials Unit at UCL: Multi-arm, multi-stage platform, umbrella and basket protocols. Clin Trials [Internet]. SAGE PublicationsSage UK: London, England; 2017 [cited 2018 Dec 9]. p. 451–61. Available from: http://journals.sagepub.com/doi/10.1177/1740774517725697
25. Porcher R, Lecocq B, Vray M, D'Andon A, Bassompierre F, Béhier J-M, et al. Adaptive Methods: When and How Should They be Used in Clinical Trials? Therapie [Internet]. 2011 [cited 2016 Jul 7];66:319–26. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21851793
26. Dimairo M, Boote J, Julious SA, Nicholl JP, Todd S. Missing steps in a staircase: a qualitative study of the perspectives of key stakeholders on the use of adaptive designs in confirmatory trials. Trials [Internet]. BioMed Central Ltd; 2015 [cited 2015 Dec 30];16:430. Available from: http://www.trialsjournal.com/content/16/1/430
27. Dimairo M, Julious SA, Todd S, Nicholl JP, Boote J. Cross-sector surveys assessing perceptions of key stakeholders towards barriers, concerns and facilitators to the appropriate use of adaptive designs in confirmatory trials. Trials [Internet]. BioMed Central Ltd; 2015 [cited 2015 Dec 27];16:585.

Available from: http://www.trialsjournal.com/content/16/1/585

28. Meurer WJ, Legocki L, Mawocha S, Frederiksen SM, Guetterman TC, Barsan W, et al. Attitudes and opinions regarding confirmatory adaptive clinical trials: a mixed methods analysis from the Adaptive Designs Accelerating Promising Trials into Treatments (ADAPT-IT) project. Trials [Internet]. BioMed Central; 2016 [cited 2016 Aug 1];17:373. Available from: http://trialsjournal.biomedcentral.com/articles/10.1186/s13063-016-1493-z
29. Morgan CC, Huyck S, Jenkins M, Chen L, Bedding a., Coffey CS, et al. Adaptive Design: Results of 2012 Survey on Perception and Use. Ther Innov Regul Sci [Internet]. 2014 [cited 2014 Nov 24];48:473–81. Available from: http://dij.sagepub.com/lookup/doi/10.1177/2168479014522468
30. Coffey CS, Levin B, Clark C, Timmerman C, Wittes J, Gilbert P, et al. Overview, hurdles, and future work in adaptive designs: perspectives from a National Institutes of Health-funded workshop. Clin Trials [Internet]. 2012 [cited 2014 Nov 24];9:671–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23250942
31. Quinlan J, Gaydos B, Maca J, Krams M. Barriers and opportunities for implementation of adaptive designs in pharmaceutical product development. Clin Trials [Internet]. 2010 [cited 2014 Nov 24];7:167–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20338900
32. Coffey CS, Kairalla JA. Adaptive clinical trials: progress and challenges. Drugs R D [Internet]. 2008 [cited 2014 Nov 24];9:229–42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18588354
33. Hartford A, Thomann M, Chen X, Miller E, Bedding A, Jorgens S, et al. Adaptive Designs: Results of 2016 Survey on Perception and Use. Ther Innov Regul Sci [Internet]. SAGE PublicationsSage CA: Los Angeles, CA; 2018 [cited 2018 Dec 10];216847901880771. Available from: http://journals.sagepub.com/doi/10.1177/2168479018807715
34. Phillips AJ, Keene ON. Adaptive designs for pivotal trials: Discussion points from the PSI Adaptive Design Expert Group. Pharm Stat [Internet]. John Wiley & Sons, Ltd.; 2006 [cited 2016 Dec 28];5:61–6. Available from: http://doi.wiley.com/10.1002/pst.206
35. Rong Y. Regulations on Adaptive Design Clinical Trials. Pharm Regul Aff Open Access [Internet]. OMICS International; 2014 [cited 2018 Mar 30];03. Available from: http://www.omicsgroup.org/journals/regulations-on-adaptive-design-clinical-trials-2167-7689.1000116.php?aid=24064
36. Park JJ, Thorlund K, Mills EJ. Critical concepts in adaptive clinical trials. Clin Epidemiol [Internet]. Dove Press; 2018 [cited 2018 Jul 12];Volume 10:343–51. Available from: https://www.dovepress.com/critical-concepts-in-adaptive-clinical-trials-peer-reviewed-article-CLEP
37. Bauer P, Brannath W. The advantages and disadvantages of adaptive designs for clinical trials. Drug Discov Today [Internet]. 2004 [cited 2014 Nov 24];9:351–7. Available from: http://www.sciencedirect.com/science/article/pii/S1359644604030235
38. Huskins WC, Fowler VG, Evans S. Adaptive Designs for Clinical Trials: Application to Healthcare Epidemiology Research. Clin Infect Dis [Internet]. 2017 [cited 2018 Jan 28]; Available from: http://academic.oup.com/cid/advance-article/doi/10.1093/cid/cix907/4600524
39. Chaitman BR, Pepine CJ, Parker JO, Skopal J, Chumakova G, Kuch J, et al. Effects of Ranolazine With Atenolol, Amlodipine, or Diltiazem on Exercise Tolerance and Angina Frequency in Patients With Severe Chronic Angina<SUBTITLE>A Randomized Controlled Trial</SUBTITLE> JAMA [Internet]. American Medical Association; 2004 [cited 2017 Jul 8];291:309. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.291.3.309
40. Zajicek JP, Hobart JC, Slade A, Barnes D, Mattison PG, MUSEC Research Group on behalf of the MR. Multiple sclerosis and extract of cannabis: results of the MUSEC trial. J Neurol Neurosurg Psychiatry [Internet]. BMJ Publishing Group Ltd; 2012 [cited 2018 Jan 23];83:1125–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22791906
41. Miller E, Gallo P, He W, Kammerman LA, Koury K, Maca J, et al. DIA's Adaptive Design Scientific Working Group (ADSWG): Best Practices Case Studies for "Less Well-understood" Adaptive Designs. Ther Innov Regul Sci [Internet]. SAGE PublicationsSage CA: Los Angeles, CA; 2017 [cited 2018 Oct 24];51:77–88. Available from: http://journals.sagepub.com/doi/10.1177/2168479016665434
42. Wang S-J, Peng H, Hung H-MJ. Evaluation of the extent of adaptation to sample size in clinical trials for cardiovascular and CNS diseases. Contemp Clin Trials [Internet]. 2018 [cited 2018 Mar 3];67:31–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29427757
43. Chen YJ, Li C, Lan KG. Sample size adjustment based on promising interim results and its application in confirmatory clinical trials. Clin Trials [Internet]. 2015 [cited 2015 Jul 21];1740774515594378-. Available from: http://ctj.sagepub.com/content/early/2015/07/17/1740774515594378.abstract?rss=1
44. Jennison C, Turnbull BW. Adaptive sample size modification in clinical trials: start small then ask for more? Stat Med [Internet]. 2015 [cited 2015 Jul 22]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/26172385
45. Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. Stat Med [Internet]. 2011 [cited 2014 Oct 31];30:3267–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22105690
46. Chiang-Stein C, Anderson K, Gallo P, Collins S. Sample size reestimation: a review and recommendations. Drug Inf J [Internet]. 2006 [cited 2014 Nov 24];40:475–84. Available from: http://cat.inist.fr/?aModele=afficheN&cpsidt=18260528
47. Friede T, Kieser M. A comparison of methods for adaptive sample size adjustment. Stat Med. 2001;20:3861–73.
48. Friede T, Kieser M. Sample size recalculation for binary data in internal pilot study designs. Pharm Stat. 2004;3:269–79.
49. Friede T, Kieser M. Sample Size Recalculation in Internal Pilot Study Designs: A Review. Biometrical J [Internet]. 2006 [cited 2015 Aug 19];48:537–55. Available from: http://doi.wiley.com/10.1002/bimj.200510238
50. Stevely A, Dimairo M, Todd S, Julious SA, Nicholl J, Hind D, et al. An Investigation of the Shortcomings of the CONSORT 2010 Statement for the Reporting of Group Sequential Randomised Controlled Trials: A Methodological Systematic Review. PLoS One [Internet]. 2015 [cited 2016 Oct 7];10:e0141104. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26528812
51. Pritchett Y, Jemiai Y, Chang Y, Bhan I, Agarwal R, Zoccali C, et al. The use of group sequential, information-based sample size re-estimation in the design of the PRIMO study of chronic kidney disease. Clin Trials [Internet]. 2011 [cited 2014 Dec 9];8:165–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21478328

52. Jennison C, Turnbull BW. Group Sequential Methods with Applications to Clinical Trials. CHAPMAN & HALL/CRC; 2000.
53. Whitehead J. The Design and Analysis of Sequential Clinical Trials. Revised 2n. Chichester, England: John Wiley & Sons Ltd; 2000.
54. Mehta CR, Tsiatis AA. Flexible Sample Size Considerations Using Information-Based Interim Monitoring. Drug Inf J [Internet]. 2001;35:1095–112. Available from: http://dij.sagepub.com/cgi/content/abstract/35/4/1095
55. Herson J, Buyse M, Wittes JT. On Stopping a Randomized Clinical Trial for Futility. In: Kowalski J, Piantadosi S, editors. Des Clin Trials Perspect Curr Issues. 1st ed. Springer; 2012. p. 109–37.
56. Gallo P, Mao L, Shih VH. Alternative Views On Setting Clinical Trial Futility Criteria. J Biopharm Stat [Internet]. 2014;24:976–93. Available from: http://www.tandfonline.com/doi/abs/10.1080/10543406.2014.932285
57. Lachin JM. Futility interim monitoring with control of type I and II error probabilities using the interim Z-value or confidence limit. Clin Trials [Internet]. 2009 [cited 2014 Nov 24];6:565–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19933716
58. Pushpakom SP, Taylor C, Kolamunnage-Dona R, Spowart C, Vora J, García-Fiñana M, et al. Telmisartan and Insulin Resistance in HIV (TAILoR): protocol for a dose-ranging phase II randomised open-labelled trial of telmisartan as a strategy for the reduction of insulin resistance in HIV-positive individuals on combination antiretroviral therapy. BMJ Open [Internet]. BMJ Publishing Group; 2015 [cited 2017 Jul 9];5:e009566. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26474943
59. Sydes MR, Parmar MKB, Mason MD, Clarke NW, Amos C, Anderson J, et al. Flexible trial design in practice - stopping arms for lack-of-benefit and adding research arms mid-trial in STAMPEDE: a multi-arm multi-stage randomized controlled trial. Trials [Internet]. Trials; 2012;13:1. Available from: http://dx.doi.org/10.1186/1745-6215-13-168
60. Parmar MKB, Barthel FM-S, Sydes M, Langley R, Kaplan R, Eisenhauer E, et al. Speeding up the evaluation of new agents in cancer. J Natl Cancer Inst [Internet]. 2008 [cited 2014 Nov 24];100:1204–14. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2528020&tool=pmcentrez&rendertype=abstract
61. Cohen DR, Todd S, Gregory WM, Brown JM. Adding a treatment arm to an ongoing clinical trial: A review of methodology and practice [Internet]. Trials. 2015 [cited 2016 Oct 7]. p. 179. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25897686
62. Magirr D, Stallard N, Jaki T. Flexible sequential designs for multi-arm clinical trials. Stat Med [Internet]. 2014 [cited 2015 Jan 8];33:3269–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24825588
63. Bretz F, Koenig F, Brannath W, Glimm E, Posch M. Adaptive designs for confirmatory clinical trials. Stat Med. 2009;28:1181–217.
64. Hommel G. Adaptive Modifications of Hypotheses After an Interim Analysis. Biometrical J [Internet]. 2001 [cited 2016 Feb 5];43:581–9. Available from: http://doi.wiley.com/10.1002/1521-4036%28200109%2943%3A5%3C581%3A%3AAID-BIMJ581%3E3.0.CO%3B2-J
65. Jaki T. Multi-arm clinical trials with treatment selection: what can be gained and at what price? Clin Investig (Lond) [Internet]. Future Science Ltd London, UK; 2015 [cited 2016 Apr 28];5:393–9. Available from: http://www.future-science.com/doi/full/10.4155/cli.15.13
66. Wason J, Magirr D, Law M, Jaki T. Some recommendations for multi-arm multi-stage trials. Stat Methods Med Res [Internet]. 2012 [cited 2018 Mar 3];25:716–27. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23242385
67. Wason J, Stallard N, Bowden J, Jennison C. A multi-stage drop-the-losers design for multi-arm clinical trials. Stat Methods Med Res [Internet]. SAGE PublicationsSage UK: London, England; 2017 [cited 2018 Mar 3];26:508–24. Available from: http://journals.sagepub.com/doi/10.1177/0962280214550759
68. Ghosh P, Liu L, Senchaudhuri P, Gao P, Mehta C. Design and monitoring of multi-arm multi-stage clinical trials. Biometrics [Internet]. Wiley/Blackwell (10.1111); 2017 [cited 2018 Jun 4];73:1289–99. Available from: http://doi.wiley.com/10.1111/biom.12687
69. Heritier S, Lô SN, Morgan CC. An adaptive confirmatory trial with interim treatment selection: Practical experiences and unbalanced randomization. Stat Med [Internet]. 2011 [cited 2015 Dec 21];30:1541–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21432887
70. Jaki T. Multi-arm clinical trials with treatment selection: what can be gained and at what price? Clin Investig (Lond) [Internet]. Future Science Ltd London, UK; 2015 [cited 2016 Dec 19];5:393–9. Available from: http://www.future-science.com/doi/10.4155/cli.15.13
71. Posch M, Koenig F, Branson M, Brannath W, Dunger-Baldauf C, Bauer P. Testing and estimation in flexible group sequential designs with adaptive treatment selection. Stat Med [Internet]. 2005 [cited 2014 Dec 6];24:3697–714. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16320264
72. Bauer P, Kieser M. Combining different phases in the development of medical treatments within a single trial. Stat Med [Internet]. 1999 [cited 2016 Mar 30];18:1833–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10407255
73. Giles FJ, Kantarjian HM, Cortes JE, Garcia-Manero G, Verstovsek S, Faderl S, et al. Adaptive Randomized Study of Idarubicin and Cytarabine Versus Troxacitabine and Cytarabine Versus Troxacitabine and Idarubicin in Untreated Patients 50 Years or Older With Adverse Karyotype Acute Myeloid Leukemia. J Clin Oncol [Internet]. American Society of Clinical Oncology; 2003 [cited 2019 Mar 21];21:1722–7. Available from: http://ascopubs.org/doi/10.1200/JCO.2003.11.016
74. Grieve AP. Response-adaptive clinical trials: case studies in the medical literature. Pharm Stat [Internet]. John Wiley & Sons, Ltd; 2017 [cited 2016 Oct 17];16:64–86. Available from: http://doi.wiley.com/10.1002/pst.1778
75. Grieve AP. Response-adaptive clinical trials: case studies in the medical literature. Pharm Stat [Internet]. Wiley-Blackwell; 2017 [cited 2018 Oct 16];16:64–86. Available from: http://doi.wiley.com/10.1002/pst.1778
76. Hu F, Rosenberger WF. The theory of response-adaptive randomization in clinical trials. New York, USA: John Wiley & Sons, Inc; 2006.
77. Nowacki AS, Zhao W, Palesch YY. A surrogate-primary replacement algorithm for response-adaptive randomization in stroke clinical trials. Stat Methods Med Res [Internet]. 2015 [cited 2015 Jan 17];1–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25586325

78. Eickhoff JC, Kim K, Beach J, Kolesar JM, Gee JR. A Bayesian adaptive design with biomarkers for targeted therapies. Clin Trials [Internet]. 2010;7:546–56. Available from: http://ctj.sagepub.com/cgi/content/abstract/7/5/546
79. Williamson SF, Jacko P, Villar SS, Jaki T. A Bayesian adaptive design for clinical trials in rare diseases. Comput Stat Data Anal. 2017;113:136–53.
80. Berry DA, Eick SG. Adaptive assignment versus balanced randomization in clinical trials: A decision analysis. Stat Med [Internet]. John Wiley & Sons, Ltd; 1995 [cited 2019 Mar 21];14:231–46. Available from: http://doi.wiley.com/10.1002/sim.4780140302
81. Chen. A seamless Phase IIB/III adaptive outcome trial: Design rationale and implementation. Clin Trials. 2014;379:393.
82. Cuffe RL, Lawrence D, Stone A, Vandemeulebroecke M. When is a seamless study desirable? Case studies from different pharmaceutical sponsors. Pharm Stat [Internet]. 2014 [cited 2014 Dec 8];13:229–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24891148
83. Donohue JF, Fogarty C, Lötvall J, Mahler DA, Worth H, Yorgancioglu A, et al. Once-daily bronchodilators for chronic obstructive pulmonary disease: indacaterol versus tiotropium. Am J Respir Crit Care Med [Internet]. 2010 [cited 2016 Mar 21];182:155–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20463178
84. Bauer P, Kieser M. Combining different phases in the development of medical treatments within a single trial. Stat Med [Internet]. 1999 [cited 2016 Dec 17];18:1833–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10407255
85. Bretz F, Schmidli H, König F, Racine A, Maurer W. Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: general concepts. Biom J [Internet]. 2006 [cited 2015 Jul 13];48:623–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16972714
86. Bauer P, Bretz F, Dragalin V, König F, Wassmer G. Twenty-five years of confirmatory adaptive designs: opportunities and pitfalls. Stat Med [Internet]. 2016 [cited 2015 Mar 18];35:325–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25778935
87. Koenig F, Brannath W, Bretz F, Posch M. Adaptive Dunnett tests for treatment selection. Stat Med [Internet]. John Wiley & Sons, Ltd; 2008 [cited 2019 Feb 15];27:1612–25. Available from: http://doi.wiley.com/10.1002/sim.3048
88. Antoniou M, Jorgensen AL, Kolamunnage-Dona R. Biomarker-Guided Adaptive Trial Designs in Phase II and Phase III: A Methodological Review. PLoS One [Internet]. Public Library of Science; 2016 [cited 2016 Feb 29];11:e0149803. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0149803
89. Liu S, Lee JJ. An overview of the design and conduct of the BATTLE trials. Chinese Clin Oncol [Internet]. 2015 [cited 2018 Jan 26];4:33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26408300
90. Barker a D, Sigman CC, Kelloff GJ, Hylton NM, Berry D a, Esserman LJ. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. Clin Pharmacol Ther. Nature Publishing Group; 2009;86:97–100.
91. Renfro LA, Mallick H, An MW, Sargent DJ, Mandrekar SJ. Clinical trial designs incorporating predictive biomarkers. Cancer Treat. Rev. 2016. p. 74–82.
92. Ondra T, Dmitrienko A, Friede T, Graf A, Miller F, Stallard N, et al. Methods for identification and confirmation of targeted subgroups in clinical trials: A systematic review. J Biopharm Stat [Internet]. Taylor & Francis; 2016 [cited 2018 Mar 3];26:99–119. Available from: http://www.tandfonline.com/doi/full/10.1080/10543406.2015.1092034
93. Chiu Y-D, Koenig F, Posch M, Jaki T. Design and estimation in clinical trials with subpopulation selection. Stat Med [Internet]. Wiley-Blackwell; 2018 [cited 2018 Oct 26]; Available from: http://doi.wiley.com/10.1002/sim.7925
94. Graf AC, Wassmer G, Friede T, Gera RG, Posch M. Robustness of testing procedures for confirmatory subpopulation analyses based on a continuous biomarker. Stat Methods Med Res [Internet]. SAGE PublicationsSage UK: London, England; 2018 [cited 2018 Jun 11];1–14. Available from: http://journals.sagepub.com/doi/10.1177/0962280218777538
95. Joshi A, Zhang J, Fang L. Statistical design for a confirmatory trial with a continuous predictive biomarker: A case study. Contemp Clin Trials [Internet]. Elsevier; 2017 [cited 2018 Oct 26];63:19–29. Available from: https://www.sciencedirect.com/science/article/pii/S1551714416304268?via%3Dihub
96. Wang SJ, Hung HMJ. Adaptive enrichment with subpopulation selection at interim: Methodologies, applications and design considerations. Contemp Clin Trials [Internet]. Elsevier B.V.; 2013;36:673–81. Available from: http://dx.doi.org/10.1016/j.cct.2013.09.008
97. Hünseler C, Balling G, Röhlig C, Blickheuser R, Trieschmann U, Lieser U, et al. Continuous Infusion of Clonidine in Ventilated Newborns and Infants. Pediatr Crit Care Med [Internet]. 2014 [cited 2018 Jan 26];15:511–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24751788
98. Hommel G. Adaptive Modifications of Hypotheses After an Interim Analysis. Biometrical J [Internet]. WILEY-VCH Verlag Berlin GmbH; 2001 [cited 2016 Dec 28];43:581–9. Available from: http://doi.wiley.com/10.1002/1521-4036%28200109%2943%3A5%3C581%3A%3AAID-BIMJ581%3E3.0.CO%3B2-J
99. Hommel G, Kropf S. Clinical Trials with an Adaptive Choice of Hypotheses. Drug Inf J [Internet]. 2001;35:1423–9. Available from: http://dij.sagepub.com/cgi/content/abstract/35/4/1423
100. Branson M, Whitehead J. Estimating a treatment effect in survival studies in which patients switch treatment. Stat Med [Internet]. 2002 [cited 2018 Oct 23];21:2449–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12205692
101. Shao J, Chang M, Chow S-C. Statistical inference for cancer trials with treatment switching. Stat Med [Internet]. Wiley-Blackwell; 2005 [cited 2018 Oct 23];24:1783–90. Available from: http://doi.wiley.com/10.1002/sim.2128
102. Skrivanek Z, Gaydos BL, Chien JY, Geiger MJ, Heathman MA, Berry S, et al. Dose-finding results in an adaptive, seamless, randomized trial of once-weekly dulaglutide combined with metformin in type 2 diabetes patients (AWARD-5). Diabetes, Obes Metab [Internet]. Blackwell Publishing Ltd; 2014 [cited 2018 Jan 26];16:748–56. Available from: http://doi.wiley.com/10.1111/dom.12305
103. Léauté-Labrèze C, Hoeger P, Mazereeuw-Hautier J, Guibaud L, Baselga E, Posiunas G, et al. A Randomized, Controlled Trial of Oral Propranolol in Infantile Hemangioma. N Engl J Med [Internet]. 2015 [cited 2015 Feb 19];372:735–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25693013
104. Mehta CR, Liu L, Theuer C. An Adaptive Population Enrichment Phase 3 Trial of TRC105 and Pazopanib Versus Pazopanib Alone in Patients with Advanced Angiosarcoma (TAPPAS Trial). Ann Oncol [Internet]. 2018 [cited 2018 Oct 28]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/30357394

105. Nogueira RG, Jadhav AP, Haussen DC, Bonafe A, Budzik RF, Bhuva P, et al. Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct. N Engl J Med [Internet]. 2017;378:11–21. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1706442

106. Collignon O, Koenig F, Koch A, Hemmings RJ, Pétavy F, Saint-Raymond A, et al. Adaptive designs in clinical trials: from scientific advice to marketing authorisation to the European Medicine Agency. Trials [Internet]. BioMed Central; 2018 [cited 2018 Nov 21];19:642. Available from: https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-018-3012-x

107. Lewis RJ, Angus DC, Laterre PF, Kjølbye AL, Van Der Meulen E, Blemings A, et al. Rationale and design of an adaptive phase 2b/3 clinical trial of selepressin for adults in septic shock: Selepressin evaluation programme for sepsis-induced shock - Adaptive clinical trial. Ann Am Thorac Soc [Internet]. American Thoracic Society; 2018 [cited 2019 Feb 3];15:250–7. Available from: http://www.atsjournals.org/doi/10.1513/AnnalsATS.201708-669SD

108. Chi L, Hung HM, Wang SJ, Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics [Internet]. 1999 [cited 2014 Nov 24];55:853–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11315017

109. Jenkins M, Stone A, Jennison C. An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. Pharm Stat [Internet]. Wiley-Blackwell; 2011 [cited 2018 Oct 28];10:347–56. Available from: http://doi.wiley.com/10.1002/pst.472

110. Wason JMS, Abraham JE, Baird RD, Gournaris I, Vallier AL, Brenton JD, et al. A Bayesian adaptive design for biomarker trials with linked treatments. Br J Cancer [Internet]. 2015 [cited 2018 Nov 17];113:699–705. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26263479

111. Saville BR, Berry SM. Efficiencies of platform clinical trials: A vision of the future. Clin Trials [Internet]. SAGE PublicationsSage UK: London, England; 2015 [cited 2018 Jun 21];13:358–66. Available from: http://journals.sagepub.com/doi/10.1177/1740774515626362

112. Bauer P, Einfalt J. Application of Adaptive Designs – a Review. Biometrical J [Internet]. 2006 [cited 2014 Nov 24];48:493–506. Available from: http://doi.wiley.com/10.1002/bimj.200510204

113. Elsäßer A, Regnstrom J, Vetter T, Koenig F, Hemmings RJ, Greco M, et al. Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency. Trials [Internet]. 2014 [cited 2014 Nov 24];15:383. Available from: http://www.trialsjournal.com/content/15/1/383

114. FDA. Guidance for Industry: Adaptive Design Clinical Trials for Drugs and Biologics [Internet]. 2010. Available from: http://www.fda.gov/downloads/Drugs/.../Guidances/ucm201790.pdf

115. CHMP. Reflection Paper on Methodological Issues in Confirmatory Clinical Trials Planned with an Adaptive Design [Internet]. 2007. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf

116. FDA. Adaptive Designs for Clinical Trials of Drugs and Biologics: Guidance for Industry [Internet]. 2018. Available from: https://www.fda.gov/downloads/drugs/guidances/ucm201790.pdf

117. Stevely A, Dimairo M, Todd S, Julious SA, Nicholl J, Hind D, et al. An Investigation of the Shortcomings of the CONSORT 2010 Statement for the Reporting of Group Sequential Randomised Controlled Trials: A Methodological Systematic Review. Shamji M, editor. PLoS One [Internet]. Public Library of Science; 2015 [cited 2015 Nov 4];10:e0141104. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0141104

118. Mistry P, Dunn JA, Marshall A. A literature review of applied adaptive design methodology within the field of oncology in randomised controlled trials and a proposed extension to the CONSORT guidelines. BMC Med Res Methodol [Internet]. BMC Medical Research Methodology; 2017;17:108. Available from: http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-017-0393-6

119. Hatfield I, Allison A, Flight L, Julious SA, Dimairo M. Adaptive designs undertaken in clinical research: a review of registered clinical trials. Trials [Internet]. BioMed Central; 2016 [cited 2016 Mar 19];17:150. Available from: http://trialsjournal.biomedcentral.com/articles/10.1186/s13063-016-1273-9

120. Sato A, Shimura M, Gosho M. Practical characteristics of adaptive design in phase 2 and 3 clinical trials. J Clin Pharm Ther [Internet]. 2017;1–11. Available from: http://doi.wiley.com/10.1111/jcpt.12617

121. Bothwell LE, Avorn J, Khan NF, Kesselheim AS. Adaptive design clinical trials: a review of the literature and ClinicalTrials.gov. BMJ Open [Internet]. British Medical Journal Publishing Group; 2018 [cited 2018 Feb 13];8:e018320. Available from: http://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2017-018320

122. Gosho M, Sato Y, Nagashima K, Takahashi S. Trends in study design and the statistical methods employed in a leading general medicine journal. J Clin Pharm Ther [Internet]. 2018 [cited 2017 Jul 30];43:36–44. Available from: http://doi.wiley.com/10.1111/jcpt.12605

123. Cerqueira FP, Jesus AMC, Cotrim MD. Adaptive Design: A Review of the Technical, Statistical, and Regulatory Aspects of Implementation in a Clinical Trial. Ther Innov Regul Sci [Internet]. SAGE PublicationsSage CA: Los Angeles, CA; 2019 [cited 2019 Mar 3];1–13. Available from: http://journals.sagepub.com/doi/10.1177/2168479019831240

124. Lin M, Lee S, Zhen B, Scott J, Horne A, Solomon G, et al. CBER's Experience With Adaptive Design Clinical Trials. Ther Innov Regul Sci [Internet]. 2015 [cited 2015 Sep 18];50:195–203. Available from: http://dij.sagepub.com/content/early/2015/09/15/2168479015604181.full?&location1=all&location2=all&row_operator2=and&term1a=adaptive+design&term_operator1=and&term_operator2=and&ct

125. Yang X, Thompson L, Chu J, Liu S, Lu H, Zhou J, et al. Adaptive Design Practice at the Center for Devices and Radiological Health (CDRH), January 2007 to May 2013. Ther Innov Regul Sci [Internet]. SAGE Publications; 2016 [cited 2016 Jul 7];50:710–7. Available from: http://dij.sagepub.com/lookup/doi/10.1177/2168479016656027

126. Dimairo M. The Utility of Adaptive Designs in Publicly Funded Confirmatory Trials [Internet]. University of Sheffield; 2016. Available from: http://etheses.whiterose.ac.uk/13981/1/Dimairo PhD Thesis 2016 WhiteRose Submission.pdf

127. Detry MA, Lewis RJ, Broglio KR, Connor JT, Berry SM, Berry DA. Standards for the Design, Conduct, and Evaluation of Adaptive Randomized Clinical Trials [Internet]. Patient-Centered Outcomes Res. Inst. 2012. Available from: http://www.pcori.org/assets/Standards-for-the-Design-Conduct-and-Evaluation-of-Adaptive-Randomized-Clinical-Trials.pdf

128. Campbell G. Similarities and Differences of Bayesian Designs and Adaptive Designs for Medical Devices: A Regulatory View. Stat Biopharm Res [Internet]. Taylor & Francis Group; 2013 [cited 2018 Jun 19];5:356–68. Available from: http://www.tandfonline.com/doi/abs/10.1080/19466315.2013.846873
129. Gaydos B, Anderson KM, Berry D, Burnham N, Chuang-Stein C, Dudinak J, et al. Good Practices for Adaptive Clinical Trials in Pharmaceutical Product Development. Ther Innov Regul Sci [Internet]. 2009;43:539–56. Available from: http://dij.sagepub.com/lookup/doi/10.1177/009286150904300503
130. Chow S-C, Chang M, Pong A. Statistical Consideration of Adaptive Methods in Clinical Development. J Biopharm Stat [Internet]. Taylor & Francis Group ; 2005 [cited 2016 Dec 18];15:575–91. Available from: http://www.tandfonline.com/doi/abs/10.1081/BIP-200062277
131. Chow S-C, Corey R. Benefits, challenges and obstacles of adaptive clinical trial designs. Orphanet J Rare Dis [Internet]. 2011 [cited 2014 Dec 3];6:79. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3248853&tool=pmcentrez&rendertype=abstract
132. Quinlan J, Krams M. Implementing adaptive designs: logistical and operational considerations. Drug Inf J [Internet]. 2006 [cited 2014 Nov 24];40:437–44. Available from: http://cat.inist.fr/?aModele=afficheN&cpsidt=18260524
133. Chow SC, Chang M, Pong A. Statistical consideration of adaptive methods in clinical development [Internet]. J. Biopharm. Stat. Taylor & Francis Group; 2005 [cited 2016 Dec 25]. p. 575–91. Available from: http://www.tandfonline.com/doi/abs/10.1081/BIP-200062277
134. Wang SJ. Perspectives on the use of adaptive designs in clinical trials. Part I. Statistical considerations and issues. J. Biopharm. Stat. 2010. p. 1090–7.
135. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med [Internet]. Public Library of Science; 2010 [cited 2016 Jun 24];7:e1000217. Available from: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000217
136. Dimairo M, Todd S, Julious S, Jaki T, Wason J, Hind D, et al. ACE Project Protocol Version 2.3: Development of a CONSORT Extension for adaptive clinical trials [Internet]. EQUATOR Netw. 2016 [cited 2018 Jan 25]. Available from: http://www.equator-network.org/wp-content/uploads/2017/12/ACE-Project-Protocol-v2.3.pdf
137. Rosenberg MJ. The agile approach to adaptive research: optimizing efficiency in clinical development. 1st ed. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2010.
138. Avery KNL, Williamson PR, Gamble C, O'Connell Francischetto E, Metcalfe C, Davidson P, et al. Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies. BMJ Open [Internet]. British Medical Journal Publishing Group; 2017 [cited 2017 Feb 20];7:e013537. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28213598
139. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for Reporting Randomized Controlled Trials in Journal and Conference Abstracts: Explanation and Elaboration. von Elm E, editor. PLoS Med [Internet]. Public Library of Science; 2008 [cited 2019 Jan 31];5:e20. Available from: https://dx.plos.org/10.1371/journal.pmed.0050020
140. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. Lancet [Internet]. 2008 [cited 2017 Aug 1];371:281–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18221781
141. Ioannidis JPA, Evans SJW, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better Reporting of Harms in Randomized Trials: An Extension of the CONSORT Statement. Ann Intern Med [Internet]. American College of Physicians; 2004 [cited 2018 Jan 13];141:781. Available from: http://annals.org/article.aspx?doi=10.7326/0003-4819-141-10-200411160-00009
142. Ton BA. Mixed Initial Boundary-Value Problems For Semi-Linear Parabolic Equations [Internet]. von Elm E, editor. Duke Math. J. Public Library of Science; 1967 [cited 2018 Jul 8]. p. 481–91. Available from: http://dx.plos.org/10.1371/journal.pmed.0050020
143. MEDLINE. Adaptive Clinical Trial MeSH Descriptor Data 2019 [Internet]. 2019 [cited 2019 Jan 6]. Available from: https://meshb.nlm.nih.gov/record/ui?ui=D000076362
144. Backonja M, Williams L, Miao X, Katz N, Chen C. Safety and efficacy of neublastin in painful lumbosacral radiculopathy: A randomized, double-blinded, placebo-controlled phase 2 trial using Bayesian adaptive design (the SPRINT trial). Pain [Internet]. Wolters Kluwer Health; 2017 [cited 2018 May 8];158:1802–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28746076
145. Barnes PJ, Pocock SJ, Magnussen H, Iqbal A, Kramer B, Higgins M, et al. Integrating indacaterol dose selection in a clinical study in COPD using an adaptive seamless design. Pulm Pharmacol Ther [Internet]. Elsevier Ltd; 2010 [cited 2014 Nov 25];23:165–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20080201
146. Jones AE, Puskarich MA, Shapiro NI, Guirgis FW, Runyon M, Adams JY, et al. Effect of Levocarnitine vs Placebo as an Adjunctive Treatment for Septic Shock. JAMA Netw Open [Internet]. American Medical Association; 2018 [cited 2019 Jan 27];1:e186076. Available from: http://jamanetworkopen.jamanetwork.com/article.aspx?doi=10.1001/jamanetworkopen.2018.6076
147. Khalil EAG, Weldegebreal T, Younis BM, Omollo R, Musa AM, Hailu W, et al. Safety and Efficacy of Single Dose versus Multiple Doses of AmBisome® for Treatment of Visceral Leishmaniasis in Eastern Africa: A Randomised Trial. Ghedin E, editor. PLoS Negl Trop Dis [Internet]. 2014 [cited 2019 Feb 7];8:21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24454970
148. Steg PG, Mehta SR, Pollack C V., Bode C, Cohen M, French WJ, et al. Anticoagulation with otamixaban and ischemic events in non-st-segment elevation acute coronary syndromes: The TAO randomized clinical trial. JAMA [Internet]. American Medical Association; 2013 [cited 2019 Feb 7];310:1145–55. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2013.277165
149. McMurray JJV, Packer M, Desai AS, Gong J, Lefkowitz MP, Rizkala AR, et al. Angiotensin–Neprilysin Inhibition versus Enalapril in Heart Failure. N Engl J Med [Internet]. Massachusetts Medical Society; 2014 [cited 2019 Mar 17];371:993–1004. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1409077
150. Gaydos B, Anderson KM, Berry D, Burnham N, Chuang-Stein C, Dudinak J, et al. Good Practices for Adaptive Clinical Trials in Pharmaceutical Product Development. Ther Innov Regul Sci [Internet]. 2009 [cited 2014 Nov 24];43:539–56. Available from: http://dij.sagepub.com/lookup/doi/10.1177/009286150904300503

151. Rosenblum M, Hanley DF. Adaptive Enrichment Designs for Stroke Clinical Trials. Stroke [Internet]. 2017 [cited 2018 Nov 1];48:2021–5. Available from: https://www.ahajournals.org/doi/10.1161/STROKEAHA.116.015342
152. Chen YJ, Li C, Lan KG. Sample size adjustment based on promising interim results and its application in confirmatory clinical trials. Clin Trials [Internet]. 2015 [cited 2015 Jul 21]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/26195615
153. Lachin JM. A review of methods for futility stopping based on conditional power. Stat Med [Internet]. 2005 [cited 2014 Nov 24];24:2747–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16134130
154. Proschan M, Lan KKG, Wittes JT. Power: Conditional, Unconditional, and Predictive. Stat Monit Clin Trials - A Unified Approach. 1st ed. New York: Springer; 2006. p. 43–66.
155. Lan KG, Simon R, Halperin M. Stochastically curtailed tests in long–term clinical trials. Seq Anal [Internet]. 1982 [cited 2014 Nov 24];37–41. Available from: http://www.tandfonline.com/doi/abs/10.1080/07474948208836014
156. Bauer P, Koenig F. The reassessment of trial perspectives from interim data—a critical view. Stat Med [Internet]. John Wiley & Sons, Ltd; 2006 [cited 2019 Feb 16];25:23–36. Available from: http://doi.wiley.com/10.1002/sim.2180
157. Herson J. Predictive probability early termination plans for phase II clinical trials. Biometrics [Internet]. 1979 [cited 2014 Nov 24];35:775–83. Available from: http://www.jstor.org/stable/2530109
158. Choi SC, Pepple PA. Monitoring Clinical Trials Based on Predictive Probability of Significance. Biometrics [Internet]. 1989 [cited 2018 Nov 1];45:317–23. Available from: https://www.jstor.org/stable/2532056?origin=crossref
159. Spiegelhalter DJ, Freedman LS, Blackburn PR. Monitoring clinical trials: Conditional or predictive power? Control Clin Trials [Internet]. 1986;7:8–17. Available from: http://linkinghub.elsevier.com/retrieve/pii/0197245686900036
160. Skrivanek Z, Berry S, Berry D, Chien J, Geiger MJ, Anderson JH, et al. Application of Adaptive Design Methodology in Development of a Long-Acting Glucagon-like Peptide-1 Analog (Dulaglutide): Statistical Design and Simulations. J Diabetes Sci Technol [Internet]. 2012;6:1305–18. Available from: http://dst.sagepub.com/cgi/content/abstract/6/6/1305
161. Ouellet D. Benefit–risk assessment: the use of clinical utility index. Expert Opin Drug Saf [Internet]. 2010 [cited 2018 Nov 1];9:289–300. Available from: https://www.tandfonline.com/doi/full/10.1517/14740330903499265
162. Thadhani R, Appelbaum E, Chang Y, Pritchett Y, Bhan I, Agarwal R, et al. Vitamin D Receptor Activation and Left Ventricular Hypertrophy in Advanced Kidney Disease. Am J Nephrol [Internet]. 2011 [cited 2018 Dec 7];33:139–49. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21242674
163. Thadhani R, Appelbaum E, Pritchett Y, Chang Y, Wenger J, Tamez H, et al. Vitamin D therapy and cardiac structure and function in patients with chronic kidney disease: the PRIMO randomized controlled trial. JAMA [Internet]. American Medical Association; 2012 [cited 2015 Oct 26];307:674–84. Available from: http://jama.jamanetwork.com/article.aspx?articleid=1104984
164. Gould AL, Shih WJ. Sample size re-estimation without unblinding for normally distributed outcomes with unknown variance. Commun Stat - Theory Methods [Internet]. Marcel Dekker, Inc.; 1992 [cited 2015 Aug 17];21:2833–53. Available from: http://www.tandfonline.com/doi/abs/10.1080/03610929208830947#.VdGRQpeTYnI
165. Gould AL. Planning and revising the sample size for a trial. Stat Med [Internet]. 1995 [cited 2015 Aug 15];14:1039-51; discussion 1053-5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7569499
166. Kieser M, Friede T. Blinded Sample Size Reestimation in Multiarmed Clinical Trials. Drug Inf J [Internet]. SAGE PublicationsSage CA: Los Angeles, CA; 2000 [cited 2019 Feb 6];34:455–60. Available from: http://journals.sagepub.com/doi/10.1177/009286150003400214
167. Posch M, Proschan MA. Unplanned adaptations before breaking the blind. Stat Med [Internet]. Wiley-Blackwell; 2012 [cited 2018 Oct 1];31:4146–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22736397
168. Chataway J, Nicholas R, Todd S, Miller DH, Parsons N, Valdés-Márquez E, et al. A novel adaptive design strategy increases the efficiency of clinical trials in secondary progressive multiple sclerosis. Mult Scler J [Internet]. 2011;17:81–8. Available from: http://msj.sagepub.com/cgi/content/abstract/17/1/81
169. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. Stat Med [Internet]. NIH Public Access; 2012 [cited 2018 Jul 12];31:2973–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22711298
170. Heatley G, Sood P, Goldstein D, Uriel N, Cleveland J, Middlebrook D, et al. Clinical Trial Design and Rationale of MOMENTUM 3: Multi-center Study of MagLev Technology in Patients Undergoing MCS Therapy With HeartMate 3^{TM} IDE Clinical Study Protocol. J Hear Lung Transplant [Internet]. 2016 [cited 2016 Feb 1];35:528–36. Available from: http://www.sciencedirect.com/science/article/pii/S1053249816000693
171. Barrington P, Chien JY, Showalter HDH, Schneck K, Cui S, Tibaldi F, et al. A 5-week study of the pharmacokinetics and pharmacodynamics of LY2189265, a novel, long-acting glucagon-like peptide-1 analogue, in patients with type 2 diabetes. Diabetes, Obes Metab [Internet]. 2011 [cited 2019 Jan 7];13:426–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21251178
172. Geiger MJ, Skrivanek Z, Gaydos B, Chien J, Berry S, Berry D, et al. An Adaptive, Dose-Finding, Seamless Phase 2/3 Study of a Long-Acting Glucagon-like Peptide-1 Analog (Dulaglutide): Trial Design and Baseline Characteristics. J Diabetes Sci Technol [Internet]. 2012;6:1319–27. Available from: http://dst.sagepub.com/cgi/content/abstract/6/6/1319
173. James ND, Sydes MR, Mason MD, Clarke NW, Anderson J, Dearnaley DP, et al. Celecoxib plus hormone therapy versus hormone therapy alone for hormone-sensitive prostate cancer: first results from the STAMPEDE multiarm, multistage, randomised controlled trial. Lancet Oncol [Internet]. 2012 [cited 2015 Jan 24];13:549–58. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3398767&tool=pmcentrez&rendertype=abstract
174. Dwan K, Kirkham JJ, Williamson PR, Gamble C. Selective reporting of outcomes in randomised controlled trials in systematic reviews of cystic fibrosis. BMJ Open [Internet]. British Medical Journal Publishing Group; 2013 [cited 2018 Oct 2];3:e002709. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23794538

175. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A-W, Cronin E, et al. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias. Siegfried N, editor. PLoS One [Internet]. Public Library of Science; 2008 [cited 2018 Oct 2];3:e3081. Available from: http://dx.plos.org/10.1371/journal.pone.0003081
176. Lancee M, Lemmens CMC, Kahn RS, Vinkers CH, Luykx JJ. Outcome reporting bias in randomized-controlled trials investigating antipsychotic drugs [Internet]. Transl. Psychiatry. Nature Publishing Group; 2017 [cited 2018 Oct 2]. p. e1232. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28895941
177. Evans S. When and How Can Endpoints Be Changed after Initiation of a Randomized Clinical Trial. PLoS Clin Trials [Internet]. Public Library of Science; 2007 [cited 2019 Mar 22];2:e18. Available from: http://dx.plos.org/10.1371/journal.pctr.0020018
178. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A-W, Cronin E, et al. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias. Siegfried N, editor. PLoS One. Public Library of Science; 2008;3:e3081.
179. Lancee M, Lemmens CMC, Kahn RS, Vinkers CH, Luykx JJ. Outcome reporting bias in randomized-controlled trials investigating antipsychotic drugs. Transl. Psychiatry. Nature Publishing Group; 2017. p. e1232.
180. Wason JMS, Mander AP, Thompson SG. Optimal multistage designs for randomised clinical trials with continuous outcomes. Stat Med [Internet]. 2012 [cited 2014 Nov 24];31:301–12. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3499690&tool=pmcentrez&rendertype=abstract
181. Wason JMS, Mander AP. Minimizing the maximum expected sample size in two-stage Phase II clinical trials with continuous outcomes. J Biopharm Stat [Internet]. Taylor & Francis; 2012 [cited 2018 Nov 25];22:836–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22651118
182. Cook JA, Julious SA, Sones W, Hampson L V, Hewitt C, Berlin JA, et al. DELTA2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. BMJ [Internet]. British Medical Journal Publishing Group; 2018 [cited 2018 Nov 5];363:k3750. Available from: http://www.bmj.com/lookup/doi/10.1136/bmj.k3750
183. Bell ML. New guidance to improve sample size calculations for trials: eliciting the target difference. Trials [Internet]. BioMed Central; 2018 [cited 2018 Nov 6];19:605. Available from: https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-018-2894-y
184. Dunnett C. Selection of the best treatment in comparison to a control with an application to a medical trial. In: Santer T, Tamhane A, editors. Des Exp Rank Sel [Internet]. New York: Marcel Dekker; 1984 [cited 2019 Feb 18]. p. 47–66. Available from: https://books.google.co.uk/books?id=1Un6FKdqUg4C&printsec=frontcover#v=onepage&q&f=false
185. Magirr D, Jaki T, Whitehead J. A generalized Dunnett test for multi-arm multi-stage clinical studies with treatment selection. Biometrika [Internet]. 2012 [cited 2015 Jan 8];99:494–501. Available from: http://biomet.oxfordjournals.org/content/99/2/494
186. Whitehead J, Jaki T. One- and two-stage design proposals for a phase II trial comparing three active treatments with control using an ordered categorical endpoint. Stat Med [Internet]. John Wiley & Sons, Ltd; 2009 [cited 2019 Jan 8];28:828–47. Available from: http://doi.wiley.com/10.1002/sim.3508
187. Jaki T, Magirr D. Designing Multi-Arm Multi-Stage Studies: R Package 'MAMS'. 2014.
188. Hwang IK, Shih WJ, De Cani JS. Group sequential designs using a family of type I error probability spending functions. Stat Med [Internet]. John Wiley & Sons, Ltd; 1990 [cited 2019 Feb 26];9:1439–45. Available from: http://doi.wiley.com/10.1002/sim.4780091207
189. Steg PG, Mehta SR, Pollack C V, Bode C, Gaudin C, Fanouillere K, et al. Design and rationale of the treatment of acute coronary syndromes with otamixaban trial: a double-blind triple-dummy 2-stage randomized trial comparing otamixaban to unfractionated heparin and eptifibatide in non-ST-segment elevation acute coronary syndro. Am Heart J [Internet]. Mosby, Inc.; 2012 [cited 2014 Nov 30];164:817–24.e13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23194481
190. Lorch U, O 'kane M, Taubel J. Three steps to writing adaptive study protocols in the early phase clinical development of new medicines. BMC Med Res Methodol. 2014;14:1–9.
191. Guetterman TC, Fetters MD, Legocki LJ, Mawocha S, Barsan WG, Lewis RJ, et al. Reflections on the Adaptive Designs Accelerating Promising Trials Into Treatments (ADAPT-IT) Process-Findings from a Qualitative Study. Clin Res Regul Aff [Internet]. 2015;32:121–30. Available from: http://www.tandfonline.com/doi/full/10.3109/10601333.2015.1079217%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/26622163%5Cnhttp://www.pubmedcartid=PMC4662412
192. Pocock SJ. Group Sequential Methods in the Design and Analysis of Clinical Trials. Biometrika [Internet]. 1977;64:191. Available from: http://www.jstor.org/stable/2335684?origin=crossref
193. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics [Internet]. 1979 [cited 2015 Mar 24];35:549–56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/497341
194. Gsponer T, Gerber F, Bornkamp B, Ohlssen D, Vandemeulebroecke M, Schmidli H. A practical guide to Bayesian group sequential designs. Pharm Stat [Internet]. Wiley-Blackwell; 2014 [cited 2018 Jul 12];13:71–80. Available from: http://doi.wiley.com/10.1002/pst.1593
195. Emerson S, Kittelson J, Gillen D. Frequentist evaluation of group sequential clinical trial designs. Stat Med [Internet]. 2007 [cited 2014 Nov 24];26:5047–80. Available from: http://onlinelibrary.wiley.com/doi/10.1002/sim.2901/full
196. Togo K, Iwasaki M. Optimal timing for interim analyses in clinical trials. J Biopharm Stat [Internet]. 2013 [cited 2014 Nov 24];23:1067–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23957516
197. Xi D, Gallo P, Ohlssen D. On the Optimal Timing of Futility Interim Analyses. Stat Biopharm Res [Internet]. 2017 [cited 2018 Nov 25];9:293–301. Available from: https://www.tandfonline.com/doi/full/10.1080/19466315.2017.1340906
198. Kelsen DP, Ginsberg R, Pajak TF, Sheahan DG, Gunderson L, Mortimer J, et al. Chemotherapy Followed by Surgery Compared with Surgery Alone for Localized Esophageal Cancer. N Engl J Med [Internet]. 1998 [cited 2018 Jul 26];339:1979–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9869669

199. Medical Research Council Oesophageal Cancer Working Group. Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. Lancet [Internet]. 2002 [cited 2018 Oct 28];359:1727–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12049861
200. Bauer P, Kohne K. Evaluation of experiments with adaptive interim analyses. Biometrics [Internet]. 1994 [cited 2014 Nov 24];50:1029–41. Available from: http://www.jstor.org/stable/2533441
201. Stahl M, Walz MK, Riera-Knorrenschild J, Stuschke M, Sandermann A, Bitzer M, et al. Preoperative chemotherapy versus chemoradiotherapy in locally advanced adenocarcinomas of the oesophagogastric junction (POET): Long-term results of a controlled randomised trial. Eur J Cancer [Internet]. Elsevier Ltd; 2017;81:183–90. Available from: http://dx.doi.org/10.1016/j.ejca.2017.04.027
202. Pocock SJ, Clayton TC, Stone GW. Challenging Issues in Clinical Trial Design: Part 4 of a 4-Part Series on Statistics for Clinical Trials. J Am Coll Cardiol [Internet]. Elsevier; 2015 [cited 2018 Jul 12];66:2886–98. Available from: https://www.sciencedirect.com/science/article/pii/S0735109715071995
203. Jansen JO, Pallmann P, Maclennan G, Campbell MK. Bayesian clinical trial designs: Another option for trauma trials? [Internet]. J. Trauma Acute Care Surg. 2017 [cited 2018 Sep 6]. p. 736–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28930965
204. Jiang Y, Zhao W, Durkalski-Mauldin V. Impact of adaptation algorithm, timing, and stopping boundaries on the performance of Bayesian response adaptive randomization in confirmative trials with a binary endpoint. Contemp Clin Trials [Internet]. Elsevier; 2017 [cited 2018 Jun 24];62:114–20. Available from: https://www.sciencedirect.com/science/article/pii/S155171441730191X?via%3Dihub
205. Du Y, Wang X, Jack Lee J. Simulation study for evaluating the performance of response-adaptive randomization. Contemp Clin Trials [Internet]. Elsevier Inc.; 2014 [cited 2015 Jan 1];40C:15–25. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84918529407&partnerID=tZOtx3y1
206. Chappell R, Durkalski V, Joffe S. University of Pennsylvania ninth annual conference on statistical issues in clinical trials: Where are we with adaptive clinical trial designs? (morning panel discussion). Clin Trials [Internet]. SAGE PublicationsSage UK: London, England; 2017 [cited 2017 Aug 19];14:441–50. Available from: http://journals.sagepub.com/doi/abs/10.1177/1740774517723590?journalCode=ctja
207. Brown CH, Ten Have TR, Jo B, Dagne G, Wyman PA, Muthén B, et al. Adaptive Designs for Randomized Trials in Public Health. Annu Rev Public Health [Internet]. NIH Public Access; 2009 [cited 2018 Mar 30];30:1–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19296774
208. Fleming TR, Sharples K, McCall J, Moore A, Rodgers A, Stewart R. Maintaining confidentiality of interim data to enhance trial integrity and credibility. Clin Trials [Internet]. NIH Public Access; 2008 [cited 2018 Jul 12];5:157–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18375654
209. He W, Gallo P, Miller E, Jemiai Y, Maca J, Koury K, et al. Addressing Challenges and Opportunities of "'Less Well-Understood'" Adaptive Designs. Ther Innov Regul Sci. 2017;51:60–8.
210. Husten L. Orexigen Released Interim Data Without Approval Of Trial Leaders. Forbes [Internet]. 2015 [cited 2018 Jul 26]; Available from: https://www.forbes.com/sites/larryhusten/2015/03/03/orexigen-released-interim-data-without-approval-of-trial-leaders/#74a030de4aef
211. Gallo P. Confidentiality and trial integrity issues for adaptive designs. Drug Inf J [Internet]. 2006 [cited 2014 Nov 24];40:445–50. Available from: http://cat.inist.fr/?aModele=afficheN&cpsidt=18260525
212. Chow S-C, Corey R, Lin M. On the independence of data monitoring committee in adaptive design clinical trials. J Biopharm Stat [Internet]. 2012;22:853–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22651119
213. Herson J. Coordinating Data Monitoring Committees and Adaptive Clinical Trial Designs. Drug Inf J [Internet]. 2008;42:297–301. Available from: http://dij.sagepub.com/cgi/content/abstract/42/4/297
214. Akacha M, Bretz F, Ohlssen D, Rosenkranz G, Schmidli H. Estimands and Their Role in Clinical Trials. Stat Biopharm Res [Internet]. Taylor & Francis; 2017 [cited 2017 Jun 12];9:268–71. Available from: https://www.tandfonline.com/doi/full/10.1080/19466315.2017.1302358
215. Akacha M, Bretz F, Ruberg S. Estimands in clinical trials – broadening the perspective. Stat Med [Internet]. 2017 [cited 2016 Dec 6];36:5–19. Available from: http://doi.wiley.com/10.1002/sim.7033
216. Gu X, Chen N, Wei C, Liu S, Papadimitrakopoulou VA, Herbst RS, et al. Bayesian Two-stage Biomarker-based Adaptive Design for Targeted Therapy Development. Stat Biosci [Internet]. NIH Public Access; 2016 [cited 2019 Mar 25];8:99–128. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27617040
217. Wittes J. Stopping a trial early - and then what? Clin Trials [Internet]. 2012 [cited 2014 Nov 24];9:714–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22879573
218. Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. JAMA [Internet]. American Medical Association; 2010 [cited 2014 Dec 15];303:1180–7. Available from: http://jama.jamanetwork.com/article.aspx?articleid=185591
219. Wears RL. Are We There Yet? Early Stopping in Clinical Trials. Ann Emerg Med [Internet]. 2015 [cited 2015 Jan 17];65:214–5. Available from: http://www.sciencedirect.com/science/article/pii/S0196064414016035
220. Hughes MD, Pocock SJ. Stopping rules and estimation problems in clinical trials. Stat Med [Internet]. 1988 [cited 2015 Apr 13];7:1231–42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3231947
221. Pocock SJ, Hughes MD. Practical problems in interim analyses, with particular regard to estimation. Control Clin Trials [Internet]. 1989 [cited 2015 Mar 26];10:209S–221S. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2605969
222. Walter SD, Han H, Briel M, Guyatt GH. Quantifying the bias in the estimated treatment effect in randomized trials having interim analyses and a rule for early stopping for futility. Stat Med [Internet]. 2017 [cited 2017 Feb 15];36:1506–18. Available from: http://doi.wiley.com/10.1002/sim.7242
223. Wang H, Rosner GL, Goodman SN. Quantifying over-estimation in early stopped clinical trials and the 'freezing effect' on subsequent research. Clin Trials [Internet]. 2016 [cited 2016 Jun 9];13:621–31. Available from: http://ctj.sagepub.com/content/early/2016/06/03/1740774516649595.full

224. Freidlin B, Korn EL. Stopping clinical trials early for benefit: impact on estimation. Clin Trials [Internet]. 2009 [cited 2014 Dec 15];6:119–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19342463
225. Bauer P, Koenig F, Brannatha W, Poscha M. Selection and bias-two hostile brothers. Stat Med [Internet]. John Wiley & Sons, Ltd; 2010 [cited 2019 Feb 8];29:1–13. Available from: http://doi.wiley.com/10.1002/sim.3716
226. Walter SD, Guyatt GH, Bassler D, Briel M, Ramsay T, Han HD. Randomised trials with provision for early stopping for benefit (or harm): The impact on the estimated treatment effect. Stat Med [Internet]. John Wiley & Sons, Ltd; 2019 [cited 2019 Mar 20];sim.8142. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.8142
227. Cameron C, Ewara E, Wilson FR, Varu A, Dyrda P, Hutton B, et al. The Importance of Considering Differences in Study Design in Network Meta-Analysis: An Application Using Anti-Tumor Necrosis Factor Drugs for Ulcerative Colitis. Med Decis Mak [Internet]. 2017;0272989X1771193. Available from: http://journals.sagepub.com/doi/10.1177/0272989X17711933
228. Mehta CR, Bauer P, Posch M, Brannath W. Repeated confidence intervals for adaptive group sequential trials. Stat Med [Internet]. 2007 [cited 2015 Mar 3];26:5422–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17918195
229. Brannath W, König F, Bauer P. Estimation in flexible two stage designs. Stat Med [Internet]. John Wiley & Sons, Ltd; 2006 [cited 2019 Feb 16];25:3366–81. Available from: http://doi.wiley.com/10.1002/sim.2258
230. Brannath W, Mehta CR, Posch M. Exact confidence bounds following adaptive group sequential tests. Biometrics [Internet]. 2009 [cited 2015 Apr 14];65:539–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18759843
231. Gao P, Liu L, Mehta C. Exact inference for adaptive group sequential designs. Stat Med [Internet]. 2013 [cited 2014 Nov 24];32:3991–4005. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23686358
232. Kunzmann K, Benner L, Kieser M. Point estimation in adaptive enrichment designs. Stat Med [Internet]. 2017 [cited 2017 Oct 10];36:3935–47. Available from: http://doi.wiley.com/10.1002/sim.7412
233. Jennison C, Turnbull BW. Analysis Following a Sequential Test. Gr Seq Methods with Appl to Clin Trials. Second Ed. Florida, USA: CHAPMAN & HALL/CRC; 2000. p. 171–87.
234. Heritier S, Lloyd CJ, Lô SN. Accurate *p* -values for adaptive designs with binary endpoints. Stat Med [Internet]. 2017 [cited 2017 Aug 9];36:2643–55. Available from: http://doi.wiley.com/10.1002/sim.7324
235. Simon R, Simon N. Inference for multimarker adaptive enrichment trials. Stat Med [Internet]. 2017 [cited 2017 Aug 17];36:4083–93. Available from: http://doi.wiley.com/10.1002/sim.7422
236. Kunz CU, Jaki T, Stallard N. An alternative method to analyse the biomarker-strategy design. Stat Med [Internet]. Wiley-Blackwell; 2018 [cited 2018 Sep 24]; Available from: http://doi.wiley.com/10.1002/sim.7940
237. Hack N, Brannath W. Estimation in adaptive group sequential trials. 2011;
238. Zhu L, Ni L, Yao B. Group sequential methods and software applications. Am Stat [Internet]. Taylor & Francis; 2011 [cited 2019 Feb 14];65:127–35. Available from: http://www.tandfonline.com/doi/abs/10.1198/tast.2011.10213
239. Tymofyeyev Y. A Review of Available Software and Capabilities for Adaptive Designs. Pract Considerations Adapt Trial Des Implement [Internet]. Springer, New York, NY; 2014 [cited 2019 Feb 14]. p. 139–55. Available from: http://link.springer.com/10.1007/978-1-4939-1100-4_8
240. Hack N, Brannath W, Brueckner M. AGSDest: Estimation in Adaptive Group Sequential Trials [Internet]. R; 2019 [cited 2019 Mar 24]. Available from: https://cran.r-project.org/web/packages/AGSDest/
241. Fernandes RM, van der Lee JH, Offringa M. A systematic review of the reporting of Data Monitoring Committees' roles, interim analysis and early termination in pediatric clinical trials. BMC Pediatr [Internet]. 2009 [cited 2015 Jan 6];9:77. Available from: http://www.biomedcentral.com/1471-2431/9/77
242. Choodari-Oskooei B, Parmar MKB, Royston P, Bowden J. Impact of lack-of-benefit stopping rules on treatment effect estimates of two-arm multi-stage (TAMS) trials with time to event outcome. Trials [Internet]. 2013 [cited 2014 Nov 24];14:23. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3599134&tool=pmcentrez&rendertype=abstract
243. Bratton DJ. Design issues and extensions of multi-arm multi-stage clinical trials [Internet]. University College London; 2015. Available from: http://discovery.ucl.ac.uk/1459437/
244. Wason JMS, Jaki T. Optimal design of multi-arm multi-stage trials. Stat Med [Internet]. 2012 [cited 2015 Sep 22];31:4269–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22826199
245. Wason J, Stallard N, Bowden J, Jennison C. A multi-stage drop-the-losers design for multi-arm clinical trials. Stat Methods Med Res [Internet]. 2014 [cited 2015 Jan 8];0962280214550759-. Available from: http://smm.sagepub.com/content/early/2014/09/15/0962280214550759.long
246. Lehmacher W, Wassmer G. Adaptive sample size calculations in group sequential trials. Biometrics [Internet]. 1999 [cited 2018 Jul 21];55:1286–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11315085
247. Vandemeulebroecke M. An Investigation of Two-Stage Tests. Stat Sin. 2006;16:933–51.
248. Graf AC, Bauer P, Glimm E, Koenig F. Maximum type 1 error rate inflation in multiarmed clinical trials with adaptive interim sample size modifications. Biom J [Internet]. 2014 [cited 2014 Dec 6];56:614–30. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4282114&tool=pmcentrez&rendertype=abstract
249. Posch M, Bauer P. Adaptive Two Stage Designs and the Conditional Error Function. Biometrical J [Internet]. 1999 [cited 2016 Mar 30];41:689–96. Available from: http://doi.wiley.com/10.1002/%28SICI%291521-4036%28199910%2941%3A6%3C689%3A%3AAID-BIMJ689%3E3.0.CO%3B2-P

250. Proschan MA, Hunsberger SA. Designed extension of studies based on conditional power. Biometrics [Internet]. 1995 [cited 2016 Mar 18];51:1315–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8589224
251. Brard C, Le Teuff G, Le Deley M-C, Hampson L V. Bayesian survival analysis in clinical trials: What methods are used in practice? Clin Trials [Internet]. 2016 [cited 2016 Oct 12];1740774516673362-. Available from: http://ctj.sagepub.com/content/early/2016/10/08/1740774516673362.abstract?rss=1
252. Wagenmakers E-J, Gronau QF, Stefan A, Dutilh G, Schönbrodt F. A Bayesian Perspective on the Proposed FDA Guidelines for Adaptive Clinical Trials [Internet]. 2018 [cited 2019 Feb 8]. Available from: https://www.bayesianspectacles.org/a-bayesian-perspective-on-the-proposed-fda-guidelines-for-adaptive-clinical-trials/
253. FDA. Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials [Internet]. 2010. Available from: https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf
254. Whitehead J. Overrunning and underrunning in sequential clinical trials. Control Clin Trials [Internet]. 1992 [cited 2015 Nov 25];13:106–21. Available from: http://www.sciencedirect.com/science/article/pii/019724569290017T
255. Hampson L V., Jennison C. Group sequential tests for delayed responses (with discussion). J R Stat Soc Ser B Stat Methodol. 2013;75:3–54.
256. Emerson S, Fleming T. Parameter estimation following group sequential hypothesis testing. Biometrika [Internet]. 1990 [cited 2014 Nov 24];77:875–92. Available from: http://biomet.oxfordjournals.org/content/77/4/875.short
257. Tröger W, Galun D, Reif M, Schumann A, Stanković N, Milićević M. Viscum album [L.] extract therapy in patients with locally advanced or metastatic pancreatic cancer: a randomised clinical trial on overall survival. Eur J Cancer [Internet]. 2013 [cited 2014 Nov 3];49:3788–97. Available from: http://www.sciencedirect.com/science/article/pii/S0959804913005509
258. MacArthur R, Hawkins T, Brown S, Lamarca A, Clay P, Barrett A, et al. Efficacy and safety of crofelemer for noninfectious diarrhea in HIV-seropositive individuals (ADVENT trial): a randomized, double-blind, placebo-controlled, two-stage study. HIV Clin Trials [Internet]. 2013 [cited 2018 Oct 24];14:261–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24334179
259. Brannath W, Zuber E, Branson M, Bretz F, Gallo P, Posch M, et al. Confirmatory adaptive designs with Bayesian decision tools for a targeted therapy in oncology. Stat Med [Internet]. John Wiley & Sons, Ltd; 2009 [cited 2019 Feb 8];28:1445–63. Available from: http://doi.wiley.com/10.1002/sim.3559
260. Marcus R, Eric P, Gabriel K. On closed testing procedures with special reference to ordered analysis of variance. Biometrika [Internet]. 1976 [cited 2015 Jan 14];63:655–60. Available from: http://biomet.oxfordjournals.org/content/63/3/655.short
261. Simes. An improved Bonferroni procedure for multiple tests of significance. Biometrika. 1986;1.
262. Kim ES, Herbst RS, Wistuba II, Lee JJ, Blumenschein GR, Tsao A, et al. The BATTLE trial: personalizing therapy for lung cancer. Cancer Discov [Internet]. American Association for Cancer Research; 2011 [cited 2019 Mar 19];1:44–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22586319
263. James ND, Sydes MR, Clarke NW, Mason MD, Dearnaley DP, Spears MR, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. Lancet (London, England) [Internet]. Elsevier; 2015 [cited 2016 Mar 1];387:1163–77. Available from: http://www.thelancet.com/article/S0140673615010375/fulltext
264. Berry SM, Connor JT, Lewis RJ. The platform trial: an efficient strategy for evaluating multiple treatments. JAMA [Internet]. American Medical Association; 2015 [cited 2016 Jan 29];313:1619–20. Available from: http://jama.jamanetwork.com/article.aspx?articleID=2210902
265. Saville BR, Berry SM. Efficiencies of platform clinical trials: A vision of the future. Clin Trials [Internet]. SAGE Publications; 2016 [cited 2016 Apr 21];13:358–66. Available from: http://ctj.sagepub.com/content/13/3/358.full
266. Gilson C, Chowdhury S, Parmar MKB, Sydes MR. Incorporating Biomarker Stratification into STAMPEDE: an Adaptive Multi-arm, Multi-stage Trial Platform. Clin Oncol [Internet]. W.B. Saunders; 2017 [cited 2019 Feb 26];29:778–86. Available from: https://www.sciencedirect.com/science/article/pii/S0936655517304387?via%3Dihub
267. Pemberton VL, Evans F, Gulin J, Rosenberg E, Addou E, Burns KM, et al. Performance and predictors of recruitment success in National Heart, Lung, and Blood Institute's cardiovascular clinical trials. Clin Trials [Internet]. 2018 [cited 2018 Oct 16];15:174077451879227. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30084662
268. Stahl M, Walz MK, Riera-Knorrenschild J, Stuschke M, Sandermann A, Bitzer M, et al. Preoperative chemotherapy versus chemoradiotherapy in locally advanced adenocarcinomas of the oesophagogastric junction (POET): Long-term results of a controlled randomised trial. Eur J Cancer [Internet]. 2017 [cited 2017 Aug 9];81:183–90. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0959804917309541
269. de Jong MD, Ison MG, Monto AS, Metev H, Clark C, O'Neil B, et al. Evaluation of Intravenous Peramivir for Treatment of Influenza in Hospitalized Patients. Clin Infect Dis [Internet]. 2014 [cited 2019 Feb 7];59:e172–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25115871
270. Harvey LA. Statistical testing for baseline differences between randomised groups is not meaningful. Spinal Cord [Internet]. Nature Publishing Group; 2018 [cited 2018 Nov 25];56:919–919. Available from: http://www.nature.com/articles/s41393-018-0203-y
271. Senn S. Testing for baseline balance in clinical trials. Stat Med [Internet]. 1994 [cited 2015 Nov 1];13:1715–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7997705
272. de Boer MR, Waterlander WE, Kuijper LDJ, Steenhuis IHM, Twisk JWR. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. Int J Behav Nutr Phys Act [Internet]. 2015 [cited 2015 Oct 30];12:4. Available from: http://www.ijbnpa.org/content/12/1/4
273. Altman DG. Comparability of Randomised Groups. Stat [Internet]. 1985 [cited 2018 Nov 25];34:125. Available from: https://www.jstor.org/stable/10.2307/2987510?origin=crossref
274. Koch A. Confirmatory Clinical Trials with an Adaptive Design. Biometrical J [Internet]. 2006 [cited 2016 Mar 10];48:574–85. Available from: http://doi.wiley.com/10.1002/bimj.200510239

275. Chang M, Chow S-C, Pong A. Adaptive design in clinical research: issues, opportunities, and recommendations. J Biopharm Stat [Internet]. 2006 [cited 2014 Nov 24];16:299-309; discussion 311-2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16724486
276. Gallo P, Chuang-Stein C. What should be the role of homogeneity testing in adaptive trials? Pharm Stat [Internet]. John Wiley & Sons, Ltd.; 2009 [cited 2016 Dec 19];8:1–4. Available from: http://doi.wiley.com/10.1002/pst.342
277. Friede T, Henderson R. Exploring changes in treatment effects across design stages in adaptive trials. Pharm Stat [Internet]. John Wiley & Sons, Ltd.; 2009 [cited 2016 Dec 28];8:62–72. Available from: http://doi.wiley.com/10.1002/pst.332
278. Wang S-J, Brannath W, Brückner M, Hung HMJ, Koch A. Unblinded Adaptive Statistical Information Design Based on Clinical Endpoint or Biomarker. Stat Biopharm Res [Internet]. 2013;5:293–310. Available from: http://www.tandfonline.com/doi/abs/10.1080/19466315.2013.791639
279. Parker RA. Testing for qualitative interactions between stages in an adaptive study. Stat Med [Internet]. John Wiley & Sons, Ltd.; 2010 [cited 2016 Dec 28];29:210–8. Available from: http://doi.wiley.com/10.1002/sim.3757
280. Gonnermann A, Framke T, Großhennig A, Koch A. No solution yet for combining two independent studies in the presence of heterogeneity. Stat Med [Internet]. Wiley-Blackwell; 2015 [cited 2016 Dec 28];34:2476–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26040434
281. Steg PG, Mehta SR, Pollack C V., Bode C, Gaudin C, Fanouillere K, et al. Design and rationale of the Treatment of Acute Coronary Syndromes with Otamixaban trial: A double-blind triple-dummy 2-stage randomized trial comparing otamixaban to unfractionated heparin and eptifibatide in non-ST-segment elevation acute coronary syndro. Am Heart J [Internet]. Mosby; 2012 [cited 2019 Feb 7];164:817–824.e13. Available from: https://www.sciencedirect.com/science/article/pii/S0002870312006643?via%3Dihub
282. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA [Internet]. American Medical Association; 2017 [cited 2017 Dec 20];318:2337. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2017.18556
283. DeMets DL, Cook TD, Buhr KA. Guidelines for statistical analysis plans. JAMA - J Am Med Assoc [Internet]. American Medical Association; 2017 [cited 2018 Mar 10];318:2301–3. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2017.18954
284. ICH. ICH E9: Statistical Principles for Clinical Trials [Internet]. 1998. Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf
285. Thorlund K, Haggstrom J, Park JJ, Mills EJ. Key design considerations for adaptive clinical trials: A primer for clinicians [Internet]. BMJ. British Medical Journal Publishing Group; 2018 [cited 2018 Mar 9]. p. k698. Available from: http://www.bmj.com/lookup/doi/10.1136/bmj.k698

Tables

Table 1. ACE checklist for the main report.

Section/Topic by Item No	Standard CONSORT 2010 checklist item	Extension for adaptive design randomised trials
Title and abstract		
1a	Identification as a randomised trial in the title	
1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts) [1]	Refer to the Adaptive designs CONSORT Extension for abstracts
Introduction		
Background and objectives		
2a	Scientific background and explanation of rationale	
2b	Specific objectives or hypotheses	
Methods		
Trial design		
3a	Description of trial design (such as parallel, factorial) including allocation ratio	
3b « Pre-planned adaptive design features		Type of adaptive design used, with details of the pre-planned trial adaptations and the statistical information informing the adaptations
3c « 3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	Important changes to the design or methods after trial commencement (such as eligibility criteria) outside the scope of the pre-planned adaptive design features, with reasons
Participants		
4a	Eligibility criteria for participants	
4b	Settings and locations where the data were collected	
Interventions		
5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes		
6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	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
6b	Any changes to trial outcomes after the trial commenced, with reasons	Any unplanned changes to trial outcomes after the trial commenced, with reasons
Sample size and operating characteristics		
7a	How sample size was determined	How sample size and operating characteristics were determined
7b	When applicable, explanation of any interim analyses and stopping guidelines	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
Randomisation		
Sequence generation		
8a	Method used to generate the random allocation sequence	
8b	Type of randomisation; details of any restriction (such as blocking and block size)	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
Allocation concealment mechanism		
9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation		
10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding		
11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	
11b	If relevant, description of the similarity of interventions	
11c Confidentiality and minimisation of operational bias		Measures to safeguard the confidentiality of interim information and minimise potential operational bias during the trial
Statistical methods		
12a	Statistical methods used to compare groups for primary and secondary outcomes	Statistical methods used to compare groups for primary and secondary outcomes, and any other outcomes used to make pre-planned adaptations
12b « Estimation and inference methods		For the implemented adaptive design features, statistical methods used to estimate treatment effects for key endpoints and to make inferences
12c « 12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results		
Participant flow (a diagram is strongly recommended)		
13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	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
13b	For each group, losses and exclusions after randomisation, together with reasons	
Recruitment and adaptations		
14a	Dates defining the periods of recruitment and follow-up	Dates defining the periods of recruitment and follow-up, for each group
14b	Why the trial ended or was stopped	See expanded E&E text for clarification
14c Adaptation decisions		Specify what trial adaptation decisions were made in light of the pre-planned decision-making criteria and observed accrued data
Baseline data		

15a « 15	A table showing baseline demographic and clinical characteristics for each group	See expanded E&E text for clarification
15b Similarity between stages		Summary of data to enable the assessment of similarity in the trial population between interim stages
Numbers analysed		
16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	See expanded E&E text for clarification
Outcomes and estimation		
17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	See expanded E&E text for clarification
17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
17c Interim results		Report interim results used to inform interim decision-making
Ancillary analyses		
18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	
Harms		
19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms) [2]	
Discussion		
Limitations		
20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	See expanded E&E text for clarification
Generalisability		
21	Generalisability (external validity, applicability) of the trial findings	See expanded E&E text for clarification
Interpretation		
22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
Other information		
Registration		
23	Registration number and name of trial registry	
Protocol		
24a « 24	Where the full trial protocol can be accessed	
Statistical analysis plan and other relevant trial documents		
24b		Where the full statistical analysis plan and other relevant trial documents can be accessed
Funding		
25	Sources of funding and other support (such as supply of drugs), role of funders	

Note: [1] Hopewell et al. 2008 [139,140]; 2 Ioannidis et al. 2004 [141]; "X« Y" means original item Y has been renumbered to X; "X«" means item reordering resulted in new item X replacing the number of the original item X; "see expanded E&E text for clarification" means the wording of the item remains unchanged but we clarify additional considerations for ADs in the E&E text.

Table 2. ACE checklist for abstracts.

Section/Topic	Standard checklist description	Extension for adaptive design randomised trials
Title	Identification of study as randomised	
Authors*	Contact details for the corresponding author	
Trial design	Description of the trial design (for example, parallel, cluster, non-inferiority)	Description of the trial design (for example, parallel, cluster, non-inferiority); include the word "adaptive" in the content or at least as a keyword
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	
Outcome	Clearly defined primary outcome for this report	See expanded E&E text for clarification
Randomisation	How participants were allocated to interventions	
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomised	Number of participants randomised to each group	
Recruitment	Trial status*	
Adaptation decisions made		Specify what trial adaptation decisions were made in light of the pre-planned decision-making criteria and observed accrued data
Numbers analysed	Number of participants analysed in each group	

Figures

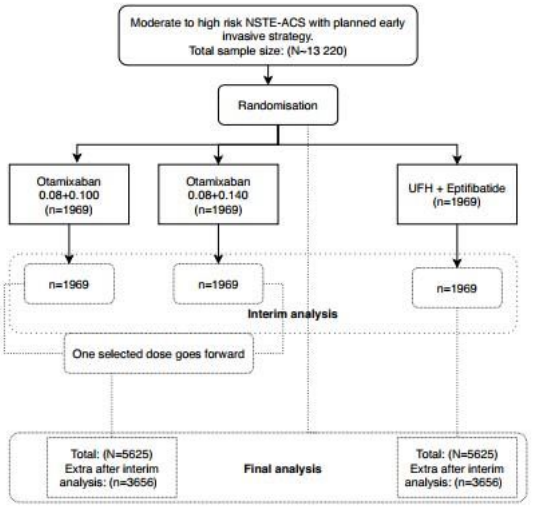

Figure 1

Sample sizes for the proposed adaptive design. Legend: Adapted from Steg et al. [197].

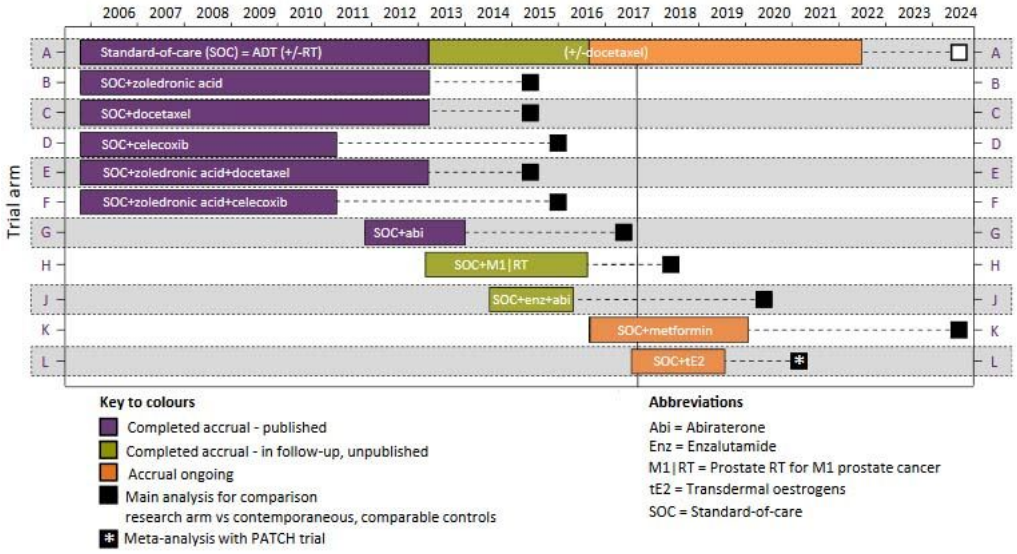

Figure 2

Recruitment and follow-up periods for STAMPEDE trial. Legend: Source – credit to Gilson et al. [267]. Reused in accordance with the terms of Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/). No changes to the original figure were made.

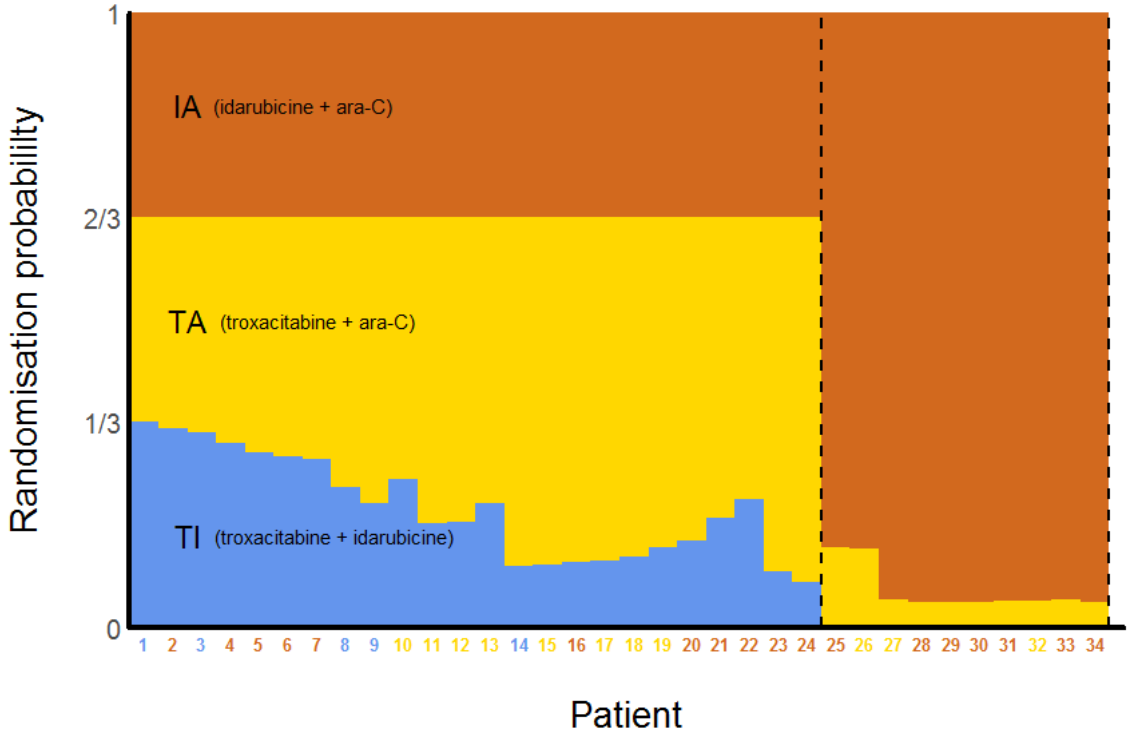

Figure 3

Change in randomisation probabilities. Legend: Source – credit to Pallmann et al [22]. Reused in accordance with the terms of Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/). No changes to the original figure were made.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplement1.docx
- supplement2.docx
- supplement3.pdf
- supplement3.pdf
- supplement5.pdf