ACE Project Protocol

Version 2.3

Trial Title	Development of a CONSORT Extension for adaptive clinical trials
Short Title	Adaptive designs CONSORT Extension (ACE) Project
Funder and Reference	NIHR CTU Support Funding and MRC HTMR
Sponsor	University of Sheffield
Version	2.3
Ethics Approval (No:)	ScHARR Research Ethics Committee (012041)
Lead Author(s)	Dimairo M
Contributors and Approvers	Steering Committee

Authors and Affiliations

Dr Munya Dimairo [1], Prof Susan Todd [2], Prof Steven Julious [1], Prof Thomas Jaki [3], Dr James Wason [4, 14], Dr Daniel Hind [1], Dr Adrian Mander [4], Prof Christopher Weir [5], Dr Franz Koenig [6], Prof Doug Altman [7], Prof Jon Nicholl [1], Prof Toshimitsu Hamasaki [8], Dr Michael Proschan [9], Dr John Scott [10], Dr Marc Walton [11], Dr Yuki Ando [12], Ms Katie Biggs [1], Dr Philip Pallmann [13], and Dr Elizabeth Coates [1]

[1] University of Sheffield, United Kingdom;
[2] University of Reading, United Kingdom;
[3] Lancaster University, United Kingdom;
[4] MRC Biostatistics Unit, University of Cambridge, United Kingdom;
[5] University of Edinburgh, United Kingdom;
[6] Medical University of Vienna, Austria;
[7] University of Oxford, United Kingdom;
[8] National Cerebral and Cardiovascular Centre, Japan;
[9] National Institute of Allergy and Infectious Diseases, National Institutes of Health, United States of America;
[10] Division of Biostatistics in the Center for Biologics Evaluation and Research, Food and Drug Administration, United States of America;
[11] Janssen Pharmaceuticals, United States of America;
[12] Pharmaceuticals and Medical Devices Agency, Japan;
[13] Cardiff University, United Kingdom;
[14] Newcastle University, United Kingdom.

Table of Contents

List of Abbreviations

ACE	Adaptive designs CONSORT Extension
ADWG	Adaptive Designs Working Group
ADSWG	Adaptive Design Scientific Working Group
CONSORT	CONsolidated Standards of Reporting Trials
CRC	Clinical Research Collaboration
CTU	Clinical Trials Unit
DFG	Deutsche Forschungsgemeinschaft (German Research Foundation)
DIA	Develop Innovate Advance
EMA	European Medicines Agency
FDA	Food and Drug Administration
EFPIA	European Federation of Pharmaceutical Industries and Associations
HRA	Health Research Authority
HTMR	Hubs for Trials Methodology Research
IBS	International Biometric Society
ICMJE	International Committee of Medical Journal Editors
ICTMC	International Clinical Trials Methodology Conference
IQR	Interquartile range
ISCB	International Society for Clinical Biostatistics
MHRA	Medicines and Healthcare products Regulatory Agency
MRC	Medical Research Council
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIAID	National Institute of Allergy and Infectious Diseases
NIH	National Institutes of Health
NIHR	National Institute for Health Research
NCVC	National Cerebral and Cardiovascular Centre
PhRMA	Pharmaceutical Research and Manufacturers of America
PMDA	Pharmaceuticals and Medical Devices Agency
PSI	Statisticians in the Pharmaceutical Industry
REC	Research Ethics Committee
ScHARR	School of Health and Related Research
SD	Standard deviation
SMG	Study Management Group

1 Introduction

1.1 Background

Adaptive designs offer pre-planned opportunities to use accumulating trial data to modify aspects of an ongoing trial while preserving its validity and integrity. Well designed and conducted adaptive designs have the potential to offer efficiency gains in addressing research objectives, as well as increased monetary and ethical advantages when investigating the benefits and risks of investigative interventions. With some adaptive designs, trials can be stopped as soon as there is sufficient evidence to answer research questions. Multiple interventions can be compared in one trial, with an option to drop futile arms early. Thus, providing considerable benefits compared to a series of independent two-arm trials (Bauer and Kieser, 1999; Bretz et al., 2006, 2009; Hommel, 2001; Jaki, 2015; Parmar et al., 2014). Such an approach accelerates the evaluation of new interventions and introduction of effective ones into practice to benefit patients, whilst reducing the burden to trial patients and saving on research resources.

Adaptive designs can also help validate trial design assumptions and modify design aspects accordingly during the trial, thereby avoiding underpowered trials when prior evidence is limited at the design stage. In addition, adaptive designs can help prospectively identify patient subgroups that are more likely to benefit from the study treatment(s). Increased uptake of adaptive designs could, therefore, increase efficiency in the delivery of research when implemented properly.

Despite the potential benefits of adaptive designs, there are multifaceted challenges hampering their routine use in clinical trials research. Transparent and adequate reporting is one of the key facilitators to mitigate some of the barriers and concerns to the use of adaptive designs, enhance their credibility in trials research, help reduce research waste, and improve reproducibility and replicability of adaptive trials. Recent research found deficiencies in the reporting of adaptive trials which may influence their credibility, usefulness to learn from and apply, and limit their ability to inform future related research (Bauer and Einfalt, 2006; Dimairo, 2016; Hatfield et al., 2016; Stevely et al., 2015). Furthermore, recent research revealed the urgent need for an adaptive designs tailored CONSORT extension (Detry, Lewis, Broglio and Connor, 2012; Dimairo, 2016; Dimairo, Julious, et al., 2015; Stevely et al., 2015).

1.2 Aims and objectives

The overall aim of this research is to develop a reporting guidance tailored to randomised trials assessing the benefits and risks of human investigative interventions using an adaptive design. Specific core objectives are to:

a) generate a list of potential items to be considered when reporting adaptive clinical trials;
b) conduct a Delphi process to survey key stakeholders in trials research, commissioning of research, approval of investigative interventions, and dissemination of research findings about their perceptions of the importance of generated reporting items and related issues they may have;
c) develop consensus on the final list of important reporting items for the development of an Adaptive designs CONSORT Extension (ACE) guidance;
d) develop the ACE guidance checklist and its explanatory supporting document; and,
e) formulate and implement dissemination strategies for these findings.

It is anticipated that the guidance document will improve transparency and adequate reporting of adaptive clinical trials, help researchers in their design, conduct and analysis, and enhance interpretation of findings by research consumers. The reporting guidance will be developed as a tool to be used by researchers, reviewers, and journal editors to promote best practice in the application and reporting of trials conducted using adaptive designs. Other research users or beneficiaries will include approvers of investigative interventions, health economists, and systematic reviewers of evidence. The guidance development process may also identify methodology gaps that could be a useful resource for methodologists.

2 Study management and membership

The Study Management Group (SMG) and the Steering Committee oversee the conduct of this research to its completion and the delivery of research objectives stated in Section 1.2. The Steering Committee is composed of international collaborators with overlapping diverse expertise and experiences including trialists, adaptive designs methods and their application, regulatory assessments, research commissioning, and journal editors. The committee also includes six members of the MRC Hubs for Trials Methodology Research (HTMR) Adaptive Designs Working Group (ADWG) and a representative of the CONSORT Executive Group. The 18 Steering

Committee members are: Dr Munya Dimairo (University of Sheffield, UK; MRC HTMR ADWG), Prof Susan Todd (University of Reading, UK), Prof Steven Julious (University of Sheffield, UK), Prof Thomas Jaki (Lancaster University, UK; MRC HTMR ADWG), Dr James Wason (MRC Biostatistics Unit, University of Cambridge, UK; MRC HTMR ADWG), Dr Daniel Hind (University of Sheffield, UK), Dr Adrian Mander (MRC Biostatistics Unit, University of Cambridge, UK; MRC HTMR ADWG), Prof Christopher Weir (University of Edinburgh, UK; MRC HTMR ADWG), Dr Franz Koenig (Medical University of Vienna, Austria), Prof Doug Altman (University of Oxford, UK; CONSORT Executive Group), Prof Jon Nicholl (University of Sheffield, UK), Prof Toshimitsu Hamasaki (NCVC, Japan), Dr Michael Proschan (NIAID, NIH, USA), Dr John Scott (FDA, USA), Dr Marc Walton (Janssen Pharmaceuticals, USA), Dr Yuki Ando (PMDA, Japan), Ms Katie Biggs (University of Sheffield, UK), and Dr Philip Pallmann (Lancaster University, UK ; MRC HTMR ADWG). In addition, Dr Elizabeth Coates a qualitative expert (University of Sheffield, UK) will be involved at different stages of the research when required.

In order to expedite the decision-making process, a SMG comprised of a subset of the Steering Committee will be actively involved in overseeing day-to-day study activities in consultation with the Steering Committee when the need arises. Key decisions made by the SMG will be communicated to the other Steering Committee members either via email correspondence or during scheduled meetings. Members of the SMG are Munya Dimairo, Katie Biggs, Thomas Jaki, Chris Weir, Susan Todd, Steven Julious, James Wason, Franz Koenig, Daniel Hind, Adrian Mander, Jon Nicholl, and Philip Pallmann.

3 Methodology

The work of the working group will: adopt a framework guiding the development of healthcare reporting guidelines (Moher et al., 2010); use published related research (Eldridge et al., 2016; Kirkham et al., 2016; Stevens et al., 2016); and specific methods highlighted in relevant proceeding sections, such as the Delphi process (Hasson et al., 2000).

3.1 The need for an adaptive designs reporting guidance and quality of reporting

Robust research on obstacles hampering the routine use of adaptive designs in clinical trials research in early and confirmatory phase across sectors has been undertaken (Coffey et al., 2012; Dimairo, Boote, et al., 2015; Dimairo, Julious, et al., 2015; Jaki, 2013; Kairalla et al., 2012; Morgan et al., 2014). Some of the leading persisting

barriers include the lack of practical knowledge and limited access to case studies of implemented adaptive designs. Hatfield et al (2016) reviewed case studies of trials implemented using adaptive designs aimed as a learning resource for trialists to help bridge the practical knowledge gap. Although the authors uncovered a number of relevant adaptive trials, they identified poor reporting of the case studies – thus hampering their usefulness as a learning resource and for their replication by other researchers.

A further review of group sequential trials with confirmatory objectives found poor reporting of items which are linked to concerns about adaptive designs raised by key stakeholders in trials research (Stevely et al., 2015). Such concerns include credibility of findings from adaptive designs in decision making to change clinical practice (Dimairo, Julious, et al., 2015). For instance, methods used to minimise operational bias due to the knowledge or leakage of interim results and to correct for potential statistical bias where appropriate were rarely disclosed.

The scope of adaptive designs (nature and extent of adaptations) employed was poorly reported and difficult to understand in reviewed case studies (Hatfield et al., 2016). Equally important, the search for adaptive designs in medical journals and clinical trials registers was problematic due to poor indexing and inadequate description of the methods (Dimairo, 2016; Hatfield et al., 2016; Yang et al., 2016).

Cross-industry key stakeholders highlighted the need and importance of an adaptive designs reporting guidance (Dimairo, Boote, et al., 2015). In a follow-up survey, the need for an adaptive designs tailored CONSORT extension gathered overwhelming support from 92% of UK Clinical Trials Units (CTUs), 88% of public funders, and 100% of private sector organisations surveyed (Dimairo, 2016; Dimairo, Julious, et al., 2015). Some studies also supported the need for better reporting on adaptive designs (Bauer and Einfalt, 2006; Yang et al., 2016). As of the 10th October 2016, there was no existing reporting guidance or related guidance under development on adaptive designs on the EQUATOR Network database based on a scoping free text search using the term “adaptive”.

In summary, the observed reporting deficiencies raise concerns about the credibility of some adaptive designs in decision-making, hamper the ability of trialists to learn about the practical application of adaptive designs, and prevent other researchers from replicating implemented adaptive trials. Following from this, a guiding belief for this project is thus that the potential efficiencies and benefits of adaptive designs in trials research can be maximised and some obstacles to their routine use mitigated through adequate transparent reporting.

3.2 Review of the literature

This work is building on an NIHR fellowship research (DRF-2012-05-182) aimed to improve the appropriate use and reporting of adaptive designs in clinical trials research (Dimairo, 2016) and other related research (Bauer and Einfalt, 2006; Detry, Lewis, Broglio, Connor, et al., 2012; Elsäßer et al., 2014; He et al., 2016; Lin et al., 2015) . With this in mind, the working group will undertake a scoping review in order to:

1) identify any form of recommendations or guidance on best practice relating to the conduct, reporting, and interpretation of adaptive trials;
2) identify potential sources of bias arising from using an adaptive design and how they could be mitigated by adequate reporting; general and design specific aspects;
3) identify key definitions of terms relating to this guideline;
4) inform the initial design of a two-stage Delphi process as a starting point for broader Steering Committee discussions to list potential reporting items (Section 3.3).

The scoping literature search will be based on a collection of grey and known related literature and a restricted MEDLINE database search. Restricted MEDLINE search will use the terms: ("*adaptive design*" OR *adaptive trial" OR "adaptive clinical trials" OR "adaptive interim" OR "flexible design")* AND (*reporting* OR *recommendation** OR (*best practice*) OR (*good practice*) OR (*panel discussion**) OR *guidance* OR *guideline** OR (*expert opinion**) OR *interpretation* OR *bias*). Most relevant paper(s) will be used to search for most similar articles using the MEDLINE similarity algorithm. Retrieved articles will be reviewed and themes generated as guided by the objectives stated above.

3.3 Stage 1: Generation of potential checklist items

This stage aims to generate a comprehensive list of reporting items for the Delphi surveys to explore the perceptions of diverse key stakeholders about their importance. Building on the scoping review findings in Section 3.2, the Steering Committee will convene a face-to-face meeting to discuss definitions of terms, concerns and how they are alleviated by better reporting, approaches to structuring the guidance document, and to draw a comprehensive list of potentially important reporting items. Small group and whole group structured discussions will be facilitated to help elicit opinion and generate discussion among members to inform the generation of a list of potential reporting items to design the round 1 Delphi survey. Generated items will be presented in logical order

to match the ordering of the existing CONSORT 2010 checklist from title to other transparency information (Schulz et al., 2010). The meeting will be audio recorded using an encrypted digital recorder and transcribed to ensure that the discussions are accurately captured.

The preliminary list will be shared with a selected external expert panel (less than 5 members) for their comments and suggestions as part of external quality control. These members will be purposively selected based on their adaptive designs expertise and approached by the SMG. The Steering Committee will review the feedback and make decisions on whether to amend the list of preliminary reporting items for the surveys or keep the existing list.

3.4 Stage 2: Delphi process

The objective of the Delphi process is to explore the perceptions about the importance of generated reporting items from a wider community of key stakeholders in clinical trials research which are described in Section 3.4.1. In addition, opinions relating to the reporting of adaptive designs will be gathered. The findings from the Delphi process will help to check for consistency in opinions between survey rounds and draw consensus on reporting items to be included in the final reporting guidance. The design of the Delphi process is based on related existing methodological guidance (Diamond et al., 2014; von der Gracht, 2012; Hasson et al., 2000). The process involves sending sequential surveys in an iterative process with Steering Committee discussions between the rounds to make related decisions. For each round, invited participants will be given up to five weeks to complete the survey. Up to six reminders will be sent out to non-responders encouraging them to complete the survey.

3.4.1 Participants selection process and how they will be approached

In the context of this study, the term key stakeholder refers to a cross-sector participant (both industry and public sector) who falls into at least one of the following categories:

a) clinical trials researchers who have used adaptive designs, have some knowledge and interest in using adaptive designs or developed adaptive design methods. These include clinical investigators, trials statisticians, trial methodologists, and health economists;

b) assessors and approvers of investigative interventions such as regulatory assessors and ethics committees;

c) beneficiaries or users of the resultant CONSORT guidance such as journal editors;

d) commissioners of research grants such as funders;

e) consumers of research results from adaptive designs and assessors of quality of evidence from trials that used adaptive designs such as systematic reviewers.

Table 1 summarises platforms to be used to reach out to potential participants to complete Delphi surveys.

Table 1: Key stakeholders and their contact platforms

Stakeholders	Platforms	Contact approach
Clinical trials researchers	• MRC HTMR (UK) • UK Clinical Research Collaboration (CRC) Network of Registered CTUs • Develop Innovate Advance (DIA) Adaptive Design Scientific Working Group (ADSWG) (USA), Working Group Adaptive Designs and Multiple Testing of International Biometric Society (IBS) Regions Austria/Swiss and Germany • Targeted conferences or organisations such as Society for Clinical Trials, International Clinical Trials Methodology Conference (ICTMC), International Society for Clinical Biostatistics (ISCB), and Statisticians in the Pharmaceutical Industry (PSI) • Known list • Targeted professional social network groups • Known researchers • Publications • Sponsors from industry (via organisations such as Pharmaceutical Research and Manufacturers of America (PhRMA) in US or European Federation of Pharmaceutical Industries and Associations (EFPIA) in Europe)	• Regular newsletter • Personal contact with adaptive designs working groups • Personal contact with known researchers who have utilised adaptive designs. For instance, Hatfield et al (2016) provide a list of phases 2, 2/3, and 3 adaptive clinical trials. Ongoing NIHR Doctoral Research Fellowship will also yield useful contacts • LinkedIn groups, such as adaptive designs group and PSI • Advertising using conference email list where possible • Publicly available contacts such as from publications
Regulators	• Food and Drug Administration (FDA) • European Medicines Agency (EMA) • Medicines and Healthcare products Regulatory Agency (MHRA), • Pharmaceuticals and Medical Devices Agency (PMDA)	• Personal contact with the help of regulatory representatives of the Steering Committee
Ethics Committees	• Network of Austrian Ethics Committee • Health Research Authority (HRA) • UK CRC Network of Registered CTUs • MRC HTMR • Other publicly available lists	• Regular newsletters • Known contacts • List of NHS RECs
Journal editors	• Leading medical research journals in publishing clinical trials, and targeted journals will be informed by journal where most adaptive designs have been published	• Personal contact of journal editors with the help of Steering Committee members • International Committee of Medical Journal Editors (ICMJE) • Blind contact approach
Funders	• Funding panels such as MRC, NIHR, and NIH • Horizon 2020 in the EU • NIH in the USA • DFG clinical trials in Germany • Above platforms	• Personal contacts • Blind contact approach
Patients	• http://www.eurordis.org/ • Sources such as cancer research network	• Patient research networks

Systematic reviewers	• Cochrane network • Relevant systematic reviews • Bodies that commissions new treatments such as NICE in the UK	• Network email list and known contacts

We will also advertise the project on social media (e.g. Twitter) and networking sites (e.g. LinkedIn) and put the information sheet and a link to survey registration on the project webpage.

3.4.2 Consent and withdrawal

Participants will provide formal consent to take part in the Delphi exercise via a tick box when they register for the survey. Details of data handling and use of data will be provided and participants can withdraw at any time, though we will keep their data up to the point of withdrawal.

3.4.3 Maintaining anonymity

It is important to capture broad perceptions of the suggested potential reporting items. This will be enhanced by approaching and inviting as many key stakeholders as possible. Participants will complete online surveys confidentially and their anonymity will be maintained during the conduct of the surveys; only the stakeholder group will be known. Consent will be sought from participants if they wish to be acknowledged for their participation in research outputs after the completion of the surveys, though they will not be linked to specific findings. Literature suggests anonymity is more likely to (von der Gracht, 2012):

a) increase participation and response rates,

b) enable responders to freely express their opinions about the importance of reporting items and any related issues they may have without undue influence of dominant members or an expert panel.

3.4.4 Sample size guide

Delphi studies are often designed with variable sample sizes and no or vague justification (Diamond et al., 2014). However, this study aims to approach as many key stakeholders of interest as possible to yield enough responders to estimate proportions with reasonable degrees of precision. With this in mind, this study intends to achieve at least 100 responders for the two survey rounds (Teare et al., 2014). We also acknowledge a potential constraint on the sample size because key stakeholders with some knowledge of adaptive designs to meaningfully contribute to the survey may be limited.

3.4.5 Delphi scoring of survey questionnaire

Participants will be asked to score their opinions of potential reporting items on an importance rating scale. The rating scoring of items will be classified as follows regarding their inclusion in the new reporting guidance:

1) scores 1 to 3 'not important for inclusion',
2) scores 4 to 6 'important, but not critical for inclusion',
3) scores 7 to 9 'critical for inclusion',
4) 'unsure' or 'don't know' for participants who are unable to give their rating opinions for some reasons.

Participants will also be given an option to feedback any comments using open-ended questionnaire fields. The objective of these comments is to uncover any additional reporting items which they view as important but were missed or considered unimportant by the Steering Committee.

3.4.6 Delphi software

Software developed by the University of Oxford will be used. The software was used to run the NETS1HD Delphi process, a Delphi for NHS England's National Maternity Dashboard, and is currently running the NETS1G and ENiGMA Delphi processes.

3.4.7 Piloting the Delphi survey

Piloting aims to troubleshoot any potential problems with the survey such as wording, scoring, and logical flow prior to launching the actual survey. The survey will be piloted on a small number of participants outside the Steering Committee. For instance, 2 to 4 participants per key stakeholder category may be used. A decision to include or exclude data from the pilot phase will be made by the Steering Committee in light of the feedback received from the pilot participants. For example, pilot data will be included only if no or minor changes that are believed negligible to alter or bias the interpretation of survey findings are made. The lead investigator in consultation with the SMG will make the selection of pilot participants where appropriate.

3.4.8 Delphi round 1 and analysis approach

Following the pilot phase, the survey design will be modified where necessary depending on feedback received prior to finalisation. The finalised survey will be 'anonymously' sent out to participants to rate their perceived importance of reporting items. The overall response rate will be computed relative to the total number of participants invited, consented, and accepted to complete the survey. The number and proportion of participants

responding to categories 'not important', 'important, but not critical', 'critically important', and 'unsure' will be presented and graphed. The denominator will be the number of responders and a 'missing' category created in the case of missing reporting item responses. Summary statistics will be presented by key stakeholder category and overall for each reporting item considered. In addition, for each reporting item, the median and interquartile range (IQR) or mean and standard deviation (SD) of ratings on an importance scale will be reported depending on the observed distributions and graphically presented, such as using clustered boxplots. A rating scale model for ordered responses (Andrich, 1978) may be used to rank the importance of reporting items as implemented in a related study (Dimairo, Julious, et al., 2015).

Qualitative data collected from participants via the open-ended questionnaire fields will be analysed thematically to identify any new proposed items for the checklist. The data will be 'read' *literally* (Mason, 2002) to discover any potential new reporting items and justification for their inclusion, and will allow us to identify uncertainty or complexity relating to the inclusion or exclusion of particular reporting items.

3.4.9 Steering Committee meeting after round 1

The Steering Committee will hold a teleconference meeting to discuss the results from round 1 and to review any feedback on additional suggested reporting items. Although the dropping of items after round 1 is not envisaged, the Steering Committee may decide to include additional reporting items depending on the importance of the reviewed feedback. Furthermore, the committee may also decide to increase the number of participants for certain key stakeholders depending on initial observed response rates.

3.4.10 Delphi round 2 and analysis approach

The Steering Committee will make a decision to proceed to round 2 together with any necessary amendments to the survey. Participants from round 1 will be re-approached and presented with their ratings and how they compare with summaries of rating from all responders. The Steering Committee will decide on the type of feedback to participants and its provision in a controlled manner. Participants will be asked to complete the survey again in light of the results from other responders. That is, each participant may decide whether to change their previous response or to retain their initial rating. In cases of strong deviation from the group responses, participants are allowed to provide reasons for their divergent rating. Such comments may help understand divergence of opinions. In addition, participants will also be allowed to complete qualitative feedback in form the of

open-ended responses as done for round 1. The change in participants' ratings from round 1 will be monitored at reporting item level and interest will be on participants who changed:

1) 1-6 'not important' or 'important, but not critical' in round 1 to 7-9 in round 2 'critically important',
2) 7-9 'critically important' in round 1 to 1-6 'not important' or 'important, but not critical' in round 2,
3) 1-3 'not important' in round 1 to 4-6 'important, but not critical' in round 2,
4) 4-6 'important, but not critical' in round 1 to 1-3 'not important' in round 2.

In addition to these changes in ratings, descriptive summaries of responses on reporting items will be reported as done for round 1. For each reporting item, the distribution of the changes in rating scores and proportion below 15% change will be reported.

In order to gauge the level of agreement between round 1 and round 2 ratings, the following statistics will be calculated and reported for each reporting item with associated 95% confidence intervals (Jakobsson and Westergren, 2005):

a) percentage agreement; percentage of participants with the same rating between rounds relative to the total responders to all rounds,
b) weighted Cohen's kappa coefficient using absolute error weights. The value of kappa was interpreted as (Landis and Koch, 1977): <0.00 'poor agreement', 0.0-0.20 'slight agreement', 0.21-0.40 'fair agreement', 0.41-0.60 'moderate agreement', 0.61-0.80 'substantial agreement', and 0.81-1.00 'almost agreement'.

These data will be graphically presented, for example, using forest plots. Any text feedback will be qualitatively analysed and quantitatively reported as themes as described for round 1.

3.4.11 Stopping criteria for the Delphi process

In this study, the objective of the Delphi process is to assess the stability of opinions which can be viewed as consistency in ratings of importance between rounds and not merely to reach consensus. Although two survey rounds are expected to reach stability based on recent related studies (Eldridge et al., 2016; Kirkham et al., 2015), the Steering Committee may decide to undertake a third survey round depending on the observed level of agreement in 'experts' ratings between the expected two rounds. The Steering Committee's decision to terminate

the Delphi process after round 2 will be based on a subjective assessment of whether the third round is likely to yield new valuable information or not aided by:

a) observed stability as measured by less than 15% change in the ratings between rounds and,
b) level of agreement in opinion ratings between the two rounds as assessed by the weighted Cohen's kappa coefficient.

3.5 Stage 3: Consensus meeting

This section describes the objectives of the consensus meeting, how the consensus criterion is defined, selection of participants and the overall activities.

3.5.1 Consensus meeting objectives

To reiterate, the Delphi process is aimed at assessing stability in opinion ratings and to quantify the level of agreement. In light of the results from the Delphi process, a face-to-face consensus meeting will be held with the objective of discussing and finalising reporting items that should be included in the reporting guidance. The decision to include reporting items will be guided by the a priori consensus definition stated in Section 3.5.2.

3.5.2 Consensus definition

The choice of a consensus criterion is guided by similar research developing healthcare reporting guidance (Eldridge et al., 2016; Kirkham et al., 2015). Reporting items that achieved the support of at least 70% of participants rating it as 'critically important' (scores 7 to 9) will be considered as having achieved the desired degree of consensus.

3.5.3 Selection of consensus group participants and its size

The composition of the consensus group will reflect the diversity of key stakeholders with current knowledge of adaptive designs, applied adaptive designs, and research consumers who may be affected or benefit from the resultant reporting guidance. In addition to the Steering Committee, additional participants representing key stakeholders of interest whose perceptions are viewed important as highlighted in Section 3.4.1 will be approached to take part. The Steering Committee will make decisions regarding additional participants to approach. A total of around 30 international participants are expected to contribute to the consensus meeting. Video conferencing will be available for international group members if needed.

3.5.4 Consensus activities

The Steering Committee will set the agenda and prepare material for the meeting to be shared with attendees in advance. The group will automatically recommend the inclusion of reporting items that reached consensus as defined in Section 3.5.2. Discussions will be held about reporting items that did not reach consensus. Reporting items that failed to reach consensus will be considered in turn based on the overall importance rating order from round 2. Following the discussion, consensus group members will anonymously be given an opportunity to make individual decisions about the inclusion of a specific item; 'keep', 'discard', and 'unsure or no opinion'. A decision to retain a reporting item will be based on achieving at least 50% support of group members deciding/wishing to keep the item. The group will agree on the final list of reporting items to be included in the reporting guidance. The discussions will be audio recorded using an encrypted digital recorder and transcribed to ensure that the content is captured accurately. In addition, the group will discuss the time frames to develop the reporting guidance and its explanatory document. After the meeting, the Steering Committee will produce a report which will be shared with the consensus meeting attendees for their comments. The final report will also be shared with Delphi process participants.

3.6 Stage 4: Development of a reporting guidance and explanatory support document

The Steering Committee will lead the development of the reporting guidance based on the agreed final list of reporting items. A detailed explanatory support document will be developed providing detailed rationale and evidence for the inclusion of the items with fundamental concept underpinning a number of adaptive designs with key relevant literature. Equally important, a number of adaptive designs exemplars will be used to illustrate the application of the guidance.

3.7 Stage 5: Piloting of the report guidance

The objective of piloting the checklist is to troubleshoot potential problems and to seek related feedback at early stages. The Steering Committee will discuss and develop pilot strategies. For instance, known investigators of ongoing and completed trials who used adaptive designs will be contacted to pilot the checklist and seek their feedback. The Steering Committee will discuss the feedback and make decisions on whether to incorporate the feedback in checklist revisions or not.

3.8 Dissemination plan

The Steering Committee will pursue both passive and active dissemination strategies (McCormack et al., 2013) aimed at maximising the awareness of the reporting guidance and engagement with the key stakeholders. The committee will also document arising methodological issues requiring prioritisation. Dissemination strategies include:

a) peer-reviewed publications in high impact medical journals;
b) hosting of research outputs through a number of platforms such as Sheffield CTRU and MRC HTMR;
c) publication of the reporting guidance and supporting document through the EQUATOR Network website;
d) organising a dissemination workshop involving multidisciplinary stakeholders;
e) organising a conference workshop session to raise awareness and disseminate the guidance;
f) presentations at relevant conferences such as SCT and ICTMC.

Detailed dissemination strategies will be discussed and developed during the course of the research.

4 Ethics approval

The study received a favourable ethics approval (012041) from the Research Ethics Committee (REC) of the School of Health and Related Research at the University of Sheffield. The conduct of the study will be guided by the granted ethics approval.

5 Discussion

There are potential benefits associated with the use of adaptive designs in clinical trials research. However, there are well-known obstacles that require addressing. Research has uncovered deficiencies in the reporting of adaptive designs which are obstructing efforts to bridge the practical knowledge gap, influencing concerns about the credibility of results from adaptive designs in decision making and the ability to replicate adaptive trials. Adequate and transparent reporting is one of the potential facilitators to mitigate some of the uncovered issues. However, there is no existing adaptive designs tailored CONSORT guidance to enhance the reporting of adaptive trials. This research, therefore, aims to develop an adaptive designs CONSORT extension

using a robust and iterative process involving a group of multidisciplinary international experts. It is hoped that the guidance document will go a long way towards:

a) mitigating concerns about adaptive designed trials,

b) improving the credibility of adaptive designed trials,

c) enhancing the interpretability of the findings from adaptive designed trials,

d) enhancing the usefulness of adaptive case studies in bridging the practical knowledge gap,

e) enhancing replication and reproducibility of adaptive designed trials,

f) enhancing the proper design and conduct of adaptive designed trials and,

g) reducing waste in clinical trials research

5.1 Acknowledgements

The study is co-funded funded by the NIHR through the CTU Support Funding and MRC HTMR. The funders have no involvement in the study design, collection, analysis, interpretation of findings, and reporting. However, research outputs will be published in line with the funders' publication policy requirements.

5.2 Funding and any additional support

We gratefully acknowledge the funding of this project by the NIHR through the CTU Support Funding and MRC HTMR. NHS Lothian via the Edinburgh Clinical Trials Unit also supports CJW in this work. MD and KB are partly funded by the NIHR and University of Sheffield to support this work.

5.3 Declaration of Conflict of Interest

All members declare that they have no conflict of interest to disclose.

6 References

1. Andrich, D. (1978), "A rating formulation for ordered response categories", *Psychometrika*, Springer-Verlag, Vol. 43 No. 4, pp. 561–573.
2. Bauer, P. and Einfalt, J. (2006), "Application of adaptive designs--a review.", *Biometrical Journal.*

Biometrische Zeitschrift, Vol. 48 No. 4, pp. 493–506.

3. Bauer, P. and Kieser, M. (1999), "Combining different phases in the development of medical treatments within a single trial.", *Statistics in Medicine*, Vol. 18 No. 14, pp. 1833–48.
4. Bretz, F., Koenig, F., Brannath, W., Glimm, E. and Posch, M. (2009), "Adaptive designs for confirmatory clinical trials", *Statistics in Medicine*, Vol. 28 No. 8, pp. 1181–1217.
5. Bretz, F., Schmidli, H., König, F., Racine, A. and Maurer, W. (2006), "Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: general concepts.", *Biometrical Journal. Biometrische Zeitschrift*, Vol. 48 No. 4, pp. 623–34.
6. Coffey, C.S., Levin, B., Clark, C., Timmerman, C., Wittes, J., Gilbert, P. and Harris, S. (2012), "Overview, hurdles, and future work in adaptive designs: perspectives from a National Institutes of Health-funded workshop.", *Clinical Trials (London, England)*, Vol. 9 No. 6, pp. 671–80.
7. Detry, M. a., Lewis, R.J., Broglio, K.R., Connor, J.T., Berry, S.M. and Berry, D. a. (2012), "Standards for the Design, Conduct, and Evaluation of Adaptive Randomized Clinical Trials", *Patient-Centered Outcomes Research Institute*, pp. 1–57.
8. Detry, M., Lewis, R., Broglio, K. and Connor, J. (2012), *Standards for the Design, Conduct, and Evaluation of Adaptive Randomized Clinical Trials*, available at: http://www.pcori.org/assets/Standards-for-the-Design-Conduct-and-Evaluation-of-Adaptive-Randomized-Clinical-Trials.pdf (accessed 24 November 2014).
9. Diamond, I.R., Grant, R.C., Feldman, B.M., Pencharz, P.B., Ling, S.C., Moore, A.M. and Wales, P.W. (2014), "Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies", *Journal of Clinical Epidemiology*, Vol. 67 No. 4, pp. 401–409.
10. Dimairo, M. (2016), *The Utility of Adaptive Designs in Publicly Funded Confirmatory Trials*, University of Sheffield, available at: http://etheses.whiterose.ac.uk/13981/1/Dimairo PhD Thesis 2016 WhiteRose Submission.pdf.
11. Dimairo, M., Boote, J., Julious, S.A., Nicholl, J.P. and Todd, S. (2015), "Missing steps in a staircase: a qualitative study of the perspectives of key stakeholders on the use of adaptive designs in confirmatory trials.", *Trials*, BioMed Central Ltd, Vol. 16 No. 1, p. 430.
12. Dimairo, M., Julious, S.A., Todd, S., Nicholl, J.P. and Boote, J. (2015), "Cross-sector surveys assessing

perceptions of key stakeholders towards barriers, concerns and facilitators to the appropriate use of adaptive designs in confirmatory trials.", *Trials*, BioMed Central Ltd, Vol. 16 No. 1, p. 585.

13. Eldridge, S.M., Chan, C.L., Campbell, M.J., Bond, C.M., Hopewell, S., Thabane, L. and Lancaster, G.A. (2016), "CONSORT 2010 statement: extension to randomised pilot and feasibility trials", *Pilot and Feasibility Studies*, BioMed Central, Vol. 2 No. 1, p. 64.
14. Elsäßer, A., Regnstrom, J., Vetter, T., Koenig, F., Hemmings, R.J., Greco, M., Papaluca-Amati, M., et al. (2014), "Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency.", *Trials*, Vol. 15 No. 1, p. 383.
15. von der Gracht, H.A. (2012), "Consensus measurement in Delphi studies. Review and implications for future quality assurance", *Technological Forecasting and Social Change*, Vol. 79 No. 8, pp. 1525–1536.
16. Hasson, F., Keeney, S. and McKenna, H. (2000), "Research guidelines for the Delphi survey technique.", *Journal of Advanced Nursing*, Vol. 32 No. 4, pp. 1008–15.
17. Hatfield, I., Allison, A., Flight, L., Julious, S.A. and Dimairo, M. (2016), "Adaptive designs undertaken in clinical research: a review of registered clinical trials", *Trials*, BioMed Central, Vol. 17 No. 1, p. 150.
18. He, W., Gallo, P., Miller, E., Jemiai, Y., Maca, J., Koury, K., Fan, X.F., et al. (2016), "Addressing Challenges and Opportunities of Less Well-Understood'' Adaptive Designs", *Therapeutic Innovation & Regulatory Science*, available at:https://doi.org/10.1177/2168479016663265.
19. Hommel, G. (2001), "Adaptive Modifications of Hypotheses After an Interim Analysis", *Biometrical Journal*, Vol. 43 No. 5, pp. 581–589.
20. Jaki, T. (2013), "Uptake of novel statistical methods for early-phase clinical studies in the UK public sector.", *Clinical Trials (London, England)*, Vol. 10 No. 2, pp. 344–6.
21. Jaki, T. (2015), "Multi-arm clinical trials with treatment selection: what can be gained and at what price?", *Clinical Investigation*, Future Science Ltd London, UK, Vol. 5 No. 4, pp. 393–399.
22. Kairalla, J. a, Coffey, C.S., Thomann, M. a and Muller, K.E. (2012), "Adaptive trial designs: a review of barriers and opportunities.", *Trials*, Trials, Vol. 13 No. 1, p. 145.
23. Kirkham, J.J., Gorst, S., Altman, D.G., Blazeby, J., Clarke, M., Devane, D., Gargon, E., et al. (2015), "COS-STAR: a reporting guideline for studies developing core outcome sets (protocol).", *Trials*, BioMed Central, Vol. 16 No. 1, p. 373.

24. Kirkham, J.J., Gorst, S., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., Gargon, E., et al. (2016), "Core Outcome Set–STAndards for Reporting: The COS-STAR Statement", pp. 1–11.
25. Lin, M., Lee, S., Zhen, B., Scott, J., Horne, A., Solomon, G. and Russek-Cohen, E. (2015), "CBER's Experience With Adaptive Design Clinical Trials", *Therapeutic Innovation & Regulatory Science*, p. 2168479015604181-.
26. Mason, J. (2002), *Qualitative Researching*, Sage Publications.
27. McCormack, L., Sheridan, S., Lewis, M., Boudewyns, V., Melvin, C.L., Kistler, C., Lux, L.J., et al. (2013), "Communication and dissemination strategies to facilitate the use of health-related evidence.", *Evidence Report/technology Assessment*, No. 213, pp. 1–520.
28. Moher, D., Schulz, K.F., Simera, I. and Altman, D.G. (2010), "Guidance for developers of health research reporting guidelines.", *PLoS Medicine*, Public Library of Science, Vol. 7 No. 2, p. e1000217.
29. Morgan, C.C., Huyck, S., Jenkins, M., Chen, L., Bedding, a., Coffey, C.S., Gaydos, B., et al. (2014), "Adaptive Design: Results of 2012 Survey on Perception and Use", *Therapeutic Innovation & Regulatory Science*, Vol. 48 No. 4, pp. 473–481.
30. Parmar, M.K.B., Carpenter, J. and Sydes, M.R. (2014), "More multiarm randomised trials of superiority are needed.", *Lancet*, Vol. 384 No. 9940, pp. 283–4.
31. Schulz, K.F., Altman, D.G. and Moher, D. (2010), "CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials.", *Annals of Internal Medicine*, Vol. 152 No. 11, pp. 726–32.
32. Stevely, A., Dimairo, M., Todd, S., Julious, S.A., Nicholl, J., Hind, D. and Cooper, C.L. (2015), "An Investigation of the Shortcomings of the CONSORT 2010 Statement for the Reporting of Group Sequential Randomised Controlled Trials: A Methodological Systematic Review", edited by Shamji, M.*PLOS ONE*, Public Library of Science, Vol. 10 No. 11, p. e0141104.
33. Stevens, G.A., Alkema, L., Black, R.E., Boerma, J.T., Collins, G.S., Ezzati, M., Grove, J.T., et al. (2016), "Guidelines for Accurate and Transparent Health Estimates Reporting: the GATHER statement.", *PLoS Medicine*, Public Library of Science, Vol. 13 No. 6, p. e1002056.
34. Teare, M.D., Dimairo, M., Shephard, N., Hayman, A., Whitehead, A. and Walters, S.J. (2014), "Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study.", *Trials*, Vol. 15 No. 1, p. 264.

35. Tobergte, D.R. and Curtis, S. (2013), “Statistical Methods for Assessing Agreement for Ordinal Data”, *Journal of Chemical Information and Modeling*, Vol. 53 No. 9, pp. 1689–1699.
36. Yang, X., Thompson, L., Chu, J., Liu, S., Lu, H., Zhou, J., Gomatam, S., et al. (2016), “Adaptive Design Practice at the Center for Devices and Radiological Health (CDRH), January 2007 to May 2013”, *Therapeutic Innovation & Regulatory Science*, SAGE Publications, p. 2168479016656027.