Reporting of randomised factorial trials: development of extensions to the CONSORT 2010 and SPIRIT 2013 guidance statements

Final Version 1.1 27-Sept-2021

Authors: Prof Alan Montgomery, Dr Brennan Kahan, Prof Diana Elbourne, Prof Elaine Beller, Prof Edmund Juszczak, Prof Paul Little, Prof An-Wen Chan & Dr Sophie Hall

Effective Date: 01-Sept-2021

STUDY PROTOCOL

RAFT

Reporting Factorial Trials

Reporting of randomised factorial trials: development of extensions to the CONSORT 2010 and SPIRIT 2013 guidance statements

Protocol version number and date: 1.1 24-Sept-2021

NCTU reference number	2013

Protocol development and sign off

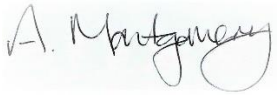
Protocol Contributors	
The following people have contributed to the development of this protocol:	
Name:	Affiliation and role:
Prof Alan Montgomery	University of Nottingham, Professor of Medical Statistics and Clinical Trials
Dr Brennan Kahan	University College London, Senior Research Fellow
Dr Sophie Hall	University of Nottingham, Research Fellow
Prof Diana Elbourne	London School of Hygiene and Tropical Medicine, Professor of Healthcare Evaluation
Ms Elaine Beller	Bond University, Associate Professor
Prof Edmund Juszczak	University of Nottingham, Professor of Clinical Trials and Statistics in Medicine
Prof Paul Little	University of Southampton, Professor of Primary Care Research
Prof An-Wen Chan	University of Toronto, Professor of Medicine

Protocol Amendments

The following amendments and/or administrative changes have been made to this protocol since the implementation of the first approved version

Amendment number	Date of amendment	Protocol version number	Type of amendment	Summary of amendment
1.1	27 Sept 2021	1.1	Non-substantial	Study duration increased from 12 to 24 months (study flow amended). Number of concept and proposed Delphi items change to 31 and 50 respectively (3.2.4. Search results and data extraction, 3.3.7 Method, and 5.1 Data description). Delphi item wording altered (Appendix 2). Participant incentives included (3.3.7. Method)

CI Signature Page	
This protocol has been approved by:	
Study Name:	RAFT
Protocol Version Number:	Version: 1.1

Protocol Version Date:	27 -Sept- 2021 (dd-mon-yyyy)
CI Name:	Professor Alan Montgomery
Study Role:	Chief Investigator
Signature and date:	
Date:	27-Sept-2021 (dd-mon-yyyy)

Administrative Information

Chief Investigator	
Prof Alan Montgomery	Professor of Medical Statistics and Clinical Trials
University of Nottingham School of Medicine, Clinical Trials Unit, Building 42, University Park, Nottingham, NG7 2RD, UK	0115 82 31612 alan.montgomery@nottingham.ac.uk

Co-Investigator (s)	
Dr Brennan Kahan	Senior Research Fellow
University College London, MRC Clinical Trials Unit, 90 High Holborn, London, WC1V 6JL, UK	b.kahan@ucl.ac.uk
Prof Diana Elbourne	Professor of Healthcare Evaluation
London School of Hygiene and Tropical Medicine, Clinical Trials Unit, Epidemiology and Population Health, Keppel Street, London, WC1E 7H, UK	diana.elbourne@lshtm.ac.uk
Prof Elaine Beller	Associate Professor
Bond University, Faculty of Health Sciences and Medicine, Level 4, Building 5, 14 University Drive, Robina, QLD 4226, Australia	ebeller@bond.edu.au
Prof Edmund Juszczak	Professor of Clinical Trials and Statistics in Medicine
University of Nottingham, School of Medicine, Clinical Trials Unit, Building 42, University Park, Nottingham, NG7 2RD, UK	ed.juszczak@nottingham.ac.uk
Prof Paul Little	Professor of Primary Care Research
University of Southampton,	psl3@soton.ac.uk

Primary Care and Population Sciences, Aldermoor Health Centre, Aldermoor Close, Southampton, SO16 5ST, UK	
Dr An-Wen Chan	Professor of Medicine
Women's College Research Institute, 76 Grenville Street, 6th Floor, Toronto, ON M5S 1B2, USA	anwen.chan@utoronto.ca
Dr Sophie Hall	Research Fellow
University of Nottingham, School of Medicine, Clinical Trials Unit, Building 42, University Park, Nottingham, NG7 2RD, UK	sophie.hall@nottingham.ac.uk

Coordinating Centre Contact Details	
Dr Sophie Hall	Research Fellow
University of Nottingham, School of Medicine, Clinical Trials Unit, Building 42, University Park, Nottingham, NG7 2RD, UK	sophie.hall@nottingham.ac.uk

STUDY SUMMARY

Title: Reporting of randomised factorial trials: development of extensions to the CONSORT 2010 and SPIRIT 2013 guidance statements.

Study Design: Delphi survey with consensus meeting.

Aim: Develop SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) and CONSORT (Consolidated Standards of Reporting Trials) extensions for factorial trials.

Participant Population and Key Eligibility Criteria: A broad range of stakeholders will be invited to participate in the study, including trialists with experience in the design/conduct/analysis/reporting of factorial trials, funding bodies, journal editors, and PPI (patient and public involvement) representatives.

Phase 1: Scoping Review: Results will be synthesised from a scoping review of relevant papers (already completed) to generate an initial 'long-list' of concepts for consideration for inclusion in the CONSORT and SPIRIT extensions.

Phase 2: Delphi Survey: An online Delphi survey will be conducted with expert stakeholders to reduce the 'long-list' to a 'short-list' of concepts.

Phase 3: Consensus Meeting: Representatives from different areas of expertise (e.g., statisticians, journal editors, trialists) will be invited to a consensus meeting to finalise the concepts to be included in the CONSORT and SPIRIT extensions for factorial trials.

Phase 4: Knowledge Translation: The concepts finalised in phase 3 will be developed into SPIRIT and CONSORT extension items for factorial trials. Guidance statements and a shared Explanation and Elaboration document will be produced.

Study Flow:

The study flow is outlined in Figure 1.

Figure 1. Study flow

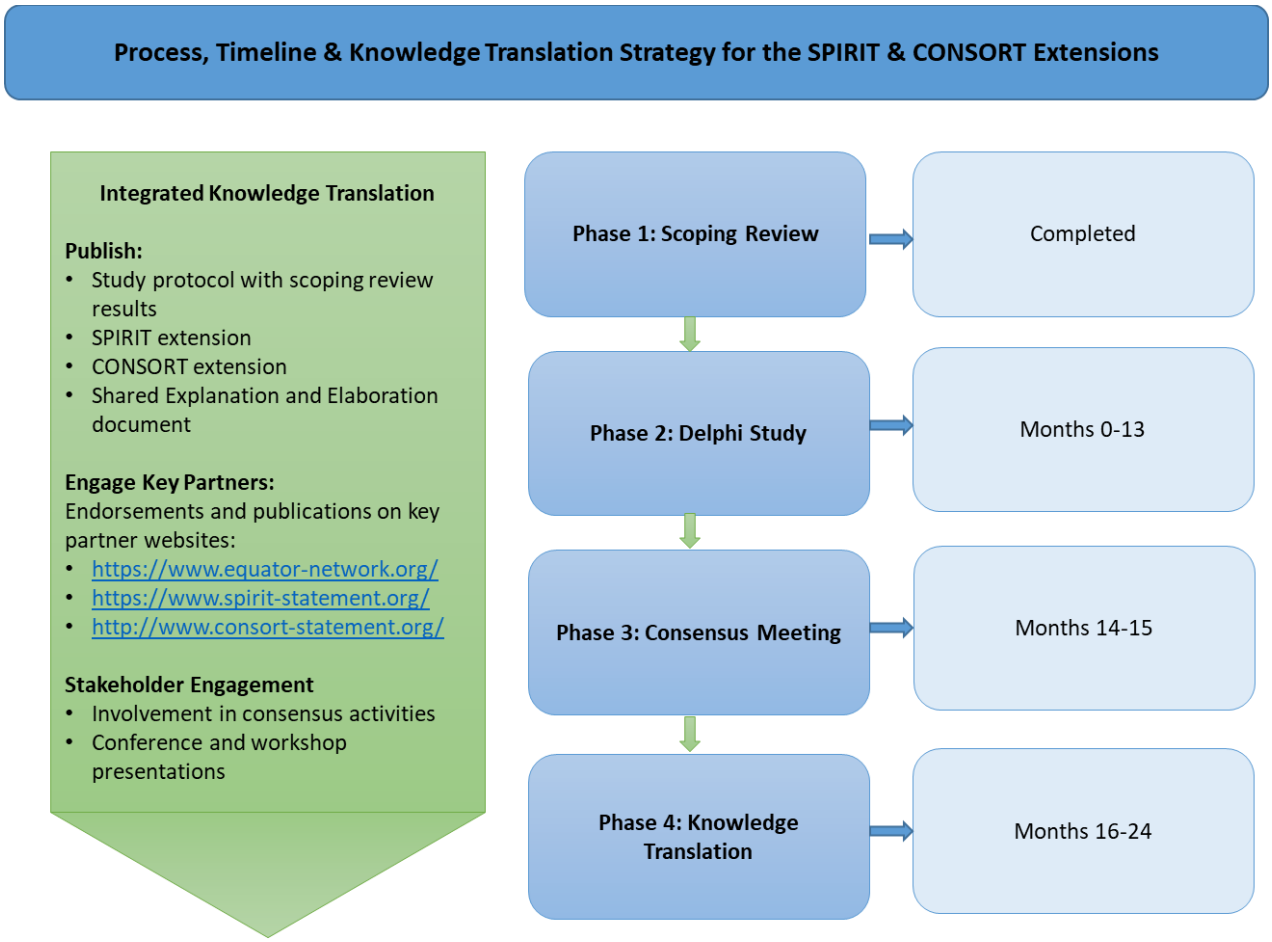


TABLE OF CONTENTS

Contents

Abbreviations and Definitions

Term	Description
CI	Chief Investigator
CONSORT	Consolidated Standards of Reporting Trials
CTU	Clinical Trials Unit
EQUATOR	Enhancing the Quality and Transparency of Health Research
HTA	Health Technology Assessment
ICH/GCP	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use/Good Clinical Practice
ISRCTN	International Standard Randomised Controlled Trial Number
MEDLINE	Medical Literature Analysis and Retrieval System Online
MESH	Medical Subject Headings
MRC	Medical Research Council
NIHR	National Institue of Health Research
PIS	Participant Information Sheet
PPI	Patient and Public Involvement
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	Prospective Register of Systematic Reviews
RAFT	Reporting Factorial Trials
SPIRIT	Standard Protocol Items
TMRP	Trials Methodology Research Partnership
UKCRC	UK Clinical Research Collaboration
UKRI	UK Research and Innovation

1. Background and Rationale

1.1. Background

Clinical trials are essential for the evaluation of healthcare interventions. However, interpretation and critical appraisal of trials is often hampered by poor reporting. The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials)[1] and CONSORT (Consolidated Standards of Reporting Trials) statements[2] are reporting guidelines which aim to facilitate better reporting. The SPIRIT statement provides recommendations for the minimum content of a trial protocol, with the aim of facilitating high quality protocols which promote rigorous trial implementation, reduce avoidable amendments, and facilitate complete appraisal of the study's scientific methods. The CONSORT Statement is a minimum set of recommendations for reporting randomised trials, with the aim of facilitating complete and transparent reporting. Both statements are widely endorsed as international standards, including by prominent medical journals, regulatory agencies, funding bodies, trial groups, and patient groups.

However, the standard guidelines do not readily apply to all trial designs and several extensions have been developed to account for this. CONSORT extensions[3] include cluster randomised[4], non-inferiority[5], within-person[6], multi-arm[7], pragmatic[8], n-of-1[9], pilot and feasibility[10] and crossover trials[11], and extensions to the SPIRIT statement include n-of-1 trials[12] and patient reported outcomes[13]. To date, there have been no extensions for factorial trials. This is an important omission since factorial trials are widely used and have distinct reporting requirements[14, 15].

Clinical trials are resource intensive, with recruitment and retention representing substantial challenges. Despite being considered the gold standard, clinical trials are prone to failure[16, 17], with a review highlighting that 44% of trials funded by the NIHR Health Technology Assessment (HTA) programme failed to reach their sample size target[18]. Trial failure not only wastes time and resources, but also poses ethical queries regarding the involvement of participants to no scientific advancement[19]. Factorial trials offer an efficient method to evaluate multiple interventions in the same trial without increasing the sample size, provided the treatments work independently (sometimes referred to as "two trials for the price of one")[15, 20, 21]. Factorial designs may also be used for other reasons (e.g., to determine whether two treatments interact); however, the focus of this study is on factorial designs which aim to efficiently assess multiple interventions in a single trial.

Use of a factorial design introduces additional methodological complexities. These are often inadequately understood, and unsurprisingly factorial trials are often poorly reported. In particular, most factorial trials are based on the assumption that the different interventions do not interact. Violation of this assumption can affect the validity of conclusions about the effects of interventions. As such, it is essential that this assumption is properly evaluated, for instance through an interaction test, and it is also recommended that that multi-arm (or ''inside-the-table'') analyses (which do not rely on the assumption of no interaction) are also presented, to check agreement with the main factorial (or "at-the-margins") analyses [14, 22].

However, in practice this is often not done. Our group recently completed a review of 100 2x2 factorial trials published between January 2015 and March 2018 and found a number of concerns around their design, analysis and reporting[14]. For instance, over one third of trials did not report any results related to the interaction. Furthermore, a number of trials used inappropriate analytical methods (i.e. basing their final analysis model on a preliminary test for interaction, which can introduce bias[14], or including a term for the interaction in the statistical model, which can alter interpretation of estimated treatment effects). Other reviews have found similar issues[15, 20, 23]. These limitations in conduct and reporting make it difficult to assess the validity of the trial conclusions, limiting their applicability for patient care and policy. It is essential that the additional complexities surrounding factorial trials are understood and accounted for in the design, conduct,

analysis, interpretation, and reporting of factorial trials. The rationale for using such a design and the assumptions made, particularly if and how an assessment of the statistical interaction was planned, must be reported transparently.

1.2. Rationale

Estimates suggest that each year there are >500 new factorial trials started[24], and 3% of all trials published in high impact medical journals employ a factorial design[25]. When correctly designed and reported, factorial trials can offer an opportunity to efficiently evaluate multiple healthcare interventions in a single trial, and there is need for specific reporting guidance to ensure the benefits offered by these trials are realised.

2. Aims, Objectives and Outcomes

2.1. Aims and Objectives

The aim of the proposed study is to improve the conduct and reporting of factorial trials by developing extensions to the CONSORT and SPIRIT statements.

Objective 1: Synthesise results from a scoping review of relevant papers (already completed) to generate an initial 'long- list' of concepts for consideration for inclusion in the CONSORT and SPIRIT extensions.

Objective 2: Conduct an online Delphi survey with expert stakeholders to reduce the 'long-list' to a 'short-list' of concepts.

Objective 3: Finalise the concepts to be included in the CONSORT and SPIRIT extensions through a consensus meeting with a broad range of stakeholders.

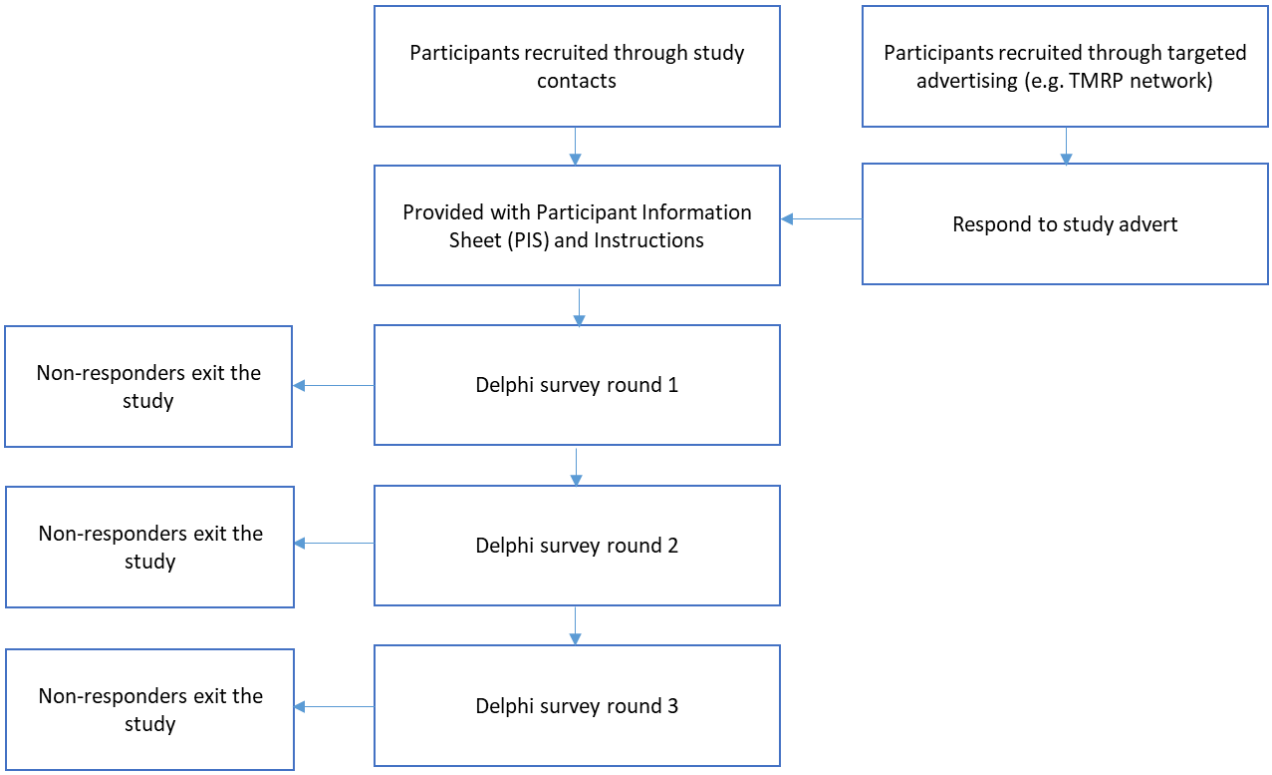
Objective 4: Develop SPIRIT and CONSORT extension items based on the concepts finalised in Objective 3 and disseminate the extensions statements along with a shared "Explanation and Elaboration" document.

3. Study Design and Setting

3.1. Study Design

This four-phase study includes a scoping review, Delphi survey, consensus meeting and knowledge translation phase. The four phases are outlined below, and participant flow for the Delphi survey is depicted in Figure 2. The study will be conducted on University of Nottingham premises, or online using Microsoft Teams, and primarily involve research staff.

Figure 2. Participant Flow for Delphi Survey



3.2. Phase One: Scoping Review

3.2.1. Objective

To generate a long list of concepts for consideration for inclusion in the CONSORT and SPIRIT extensions based on reporting or methodological considerations published in the literature.

Generation of Search Terms

MEDLINE search terms were developed by a research librarian with input from the team and guided by a recent similar review in this area[14]. Given that this was a scoping review, PROSPERO (Prospective Register of Systematic Reviews) registration was not required.

The full search strategy was: (((Factorial OR "2 x 2" OR 2x2 OR "Two by two" OR "Four arm" OR "Four-arm")) AND ("Randomized Controlled Trials as Topic"[Mesh] OR Trial OR Trials)) AND ("Research Design"[Mesh] OR "Research Design" OR "Research Designs" OR "Factorial Design" OR "Factorial Designs" OR "Research Techniques" OR "Research Technique" OR "Experimental Design" OR "Experimental Designs" OR "Research Methodology" OR "Research methods").

The above search was supplemented with two articles from our personal collections that were not currently available in the public domain at that time[7, 14].

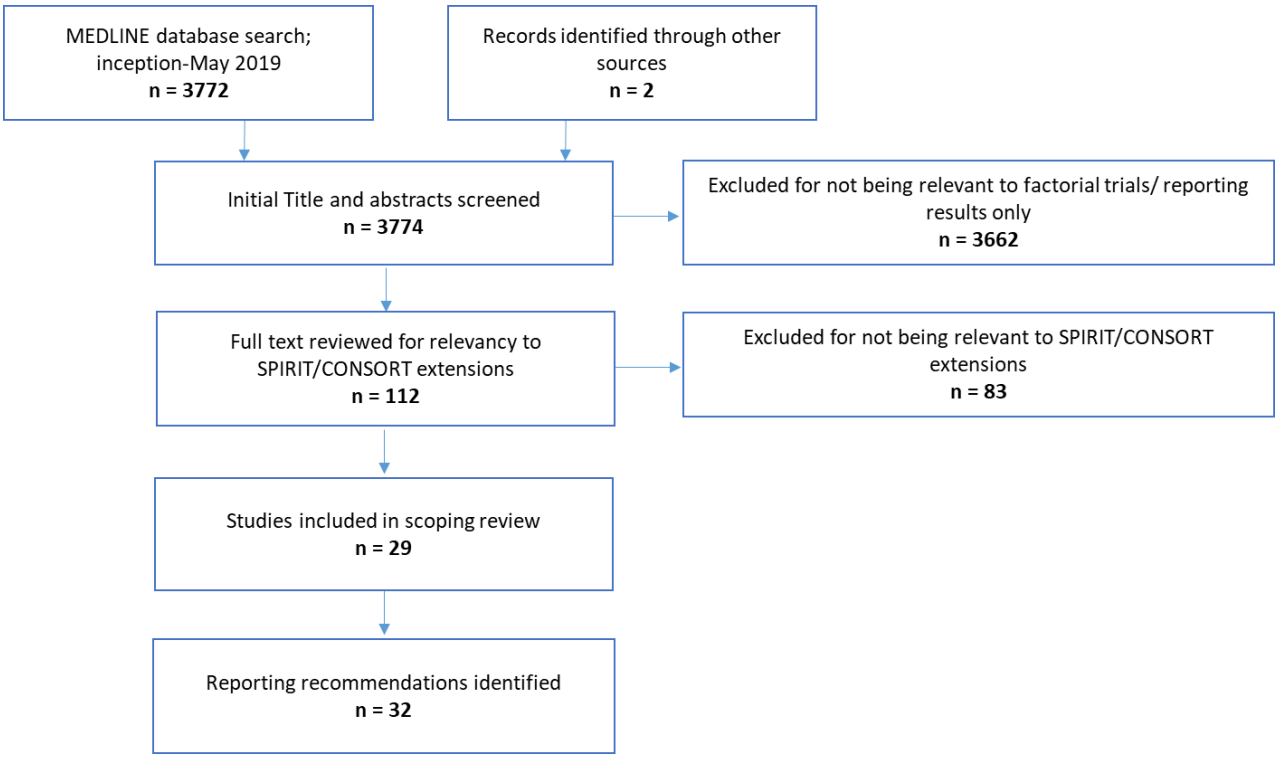
3.2.2. Database Searches

In line with previous reviews searching a similar topic, searches were conducted in MEDLINE from inception to May 2019[14].

3.2.3. Study Selection

Articles were included in the scoping review if they discussed methodological or reporting issues relevant to factorial trials. Titles and abstracts were screened by one author (EB) to remove articles that were not relevant (e.g., those that reported results of a randomised trial, but not methodological issues). Another author (BK) then further screened abstracts and full texts to assess whether articles were relevant to the SPIRIT/CONSORT extension. Reporting suggestions and methodological issues requiring additional reporting were then recorded. The records identified, included, and excluded are displayed in Figure 3.

Figure 3. PRISMA Flow Diagram for Scoping Review



3.2.4. Search Results and Data Extraction

A total of 3772 articles were identified from the Medline search, which was supplemented by two articles were included from our personal collections. After title/abstract screening, we were left with 112 possible articles. We excluded a further 83 articles after full text screening, leaving 29 included papers (see Appendix 1).

We identified 31 reporting recommendations from the 29 articles (see Appendix 2). These 31 concepts will form the basis for the questions in the Delphi survey.

3.3. Phase Two: Delphi Survey

3.3.1. Objective

Conduct an online Delphi survey with expert stakeholders to reduce the 'long-list' to a 'short-list' of concepts.

3.3.2. Participants

The sample size for a Delphi survey depends on the group dynamics for reaching consensus as opposed to statistical power, since it is a qualitative study. Therefore, guidelines vary widely from as low as 10 participants[26], or up to 50[27], with other recommendations suggesting the group characteristics are taken into consideration ensuring there are approximately 5-10 participants representing different groups[28]. Our previous experience suggests we are able to recruit around 100-130 participants[29], and this is an appropriate target sample given that we aim to recruit participants with a broad range of experience in the design/conduct/analysis/reporting of factorial trials, knowledge users, representatives from funding bodies, journal editors, and PPI (patient and public involvement) representatives.

3.3.3. Eligibility

Methodologists and trialists will be identified through the team's extensive database of international contacts, relevant publications (lead authors and methodological co-authors who have published either factorial clinical trials, or methodological work related to factorial trials), as well as through relevant networks (for example, MRC-NIHR Trials Methodology Research Partnership (TMRP), network of UKCRC-registered Clinical Trials Units). We will invite representatives from funding bodies (NIHR, UKRI, charities), and journal editors with experience in CONSORT or SPIRIT statements. We will identify experienced PPI representatives with an interest in trials and/or reporting.

3.3.4. Inclusion Criteria

- At least a basic understanding or prior experience of factorial trials (to be judged by study team)
- Ability to speak, read and write English
- Availability to respond during the specified data collection period
- Regular access to Broadband internet

3.3.5. Exclusion Criteria

- None

3.3.6. Consent

A Participant Information Sheet (PIS) will be sent during the recruitment process describing the purpose of the study, and the process of the Delphi survey. Additionally a keyword definitions document will be provided, briefly describing key terminologies that are frequently used in factorial trials. Information on data usage, sharing, anonymity and right to withdraw will be included in the information sheet. The PIS will be written using non-expert language. Delphi survey participants will be encouraged to contact the research team with any questions. The DelphiManager software includes a consent checkbox, complying with GDPR regulations.

3.3.7. Method

The amount of personally identifiable information collected will be minimised to include: background/discipline experience, type and duration of experience in factorial trials, location, and contact details to enable personalised reminders.

Results from the scoping review indicate that the number of preliminary items for the Delphi survey will be around 50 (relating to the 31 concepts, 18 apply to SPIRIT and CONSORT, 13 apply to CONSORT only). Conducting more than one Delphi survey round allows participants to suggest new items for inclusion in subsequent round(s). Therefore, we propose a 3 round survey to enable participants to suggest new items in the first round. All items will be evaluated at least twice. Items from round 1 that are not at consensus after round 2 will be evaluated three times. After round 2, items that have reached consensus will be removed before proceeding to round 3. Previous experience suggests that this process will be sufficient to achieve consensus on a short-list of essential items to be taken forward to the consensus meeting in phase 3, while reducing risk of attrition bias that is possible with more survey rounds.

The Delphi exercise will be conducted according to published guidance[30]. The Delphi survey will be facilitated using DelphiManager software[31] by the named research fellow. For each survey round, participants will be invited by email (with weekly email reminders) with up to four weeks to complete each survey round. Only those participants who completed round 1 will be invited for round 2, and only participants who completed round 2 will be invited for round 3.

The first survey round will be based on the 31 concepts identified in the scoping review. Participants will be asked to score each of the 31 concepts (leading to 50 individual items) on the importance of including it in the final set of guidelines for the SPIRIT extension and then for the CONSORT extension (as separate items). We will use a 9-point Likert scale:

- 1-3 = not important
- 4-6 = important but not critical
- 7-9 = critical

In addition, the following options will be provided:

- An ‘unable to rate option’
- Space for free text to explain a given score
- Nominate additional items for inclusion (available in round 1 only)

Participants who complete all 3-survey rounds will be offered a certificate of study completion and the opportunity to be entered into a prize draw to receive one of two £100 (GBP equivalent) online shopping vouchers.

3.3.8. Analysis

The second survey round will include all concepts from round 1, as well as new concepts that were suggested by participants in round 1. Participants will be provided with a summary of scores from other participants along with their own score. Feedback will be presented at a whole group level (i.e., not split by stakeholder group).

Each item will be tabulated and categorised as follows[32]:

- Consensus in = ≥70% participants scoring 7-9 and <15% participants scoring 1-3
- Consensus out = ≥70% participants scoring 1-3 and <15% of participants scoring 7-9
- No consensus = anything else

Concepts that do not reach consensus will form the focus of the consensus meeting.

3.4. Phase Three: Consensus Meeting

3.4.1. Objective

Organise a meeting of a broad range of stakeholders to establish consensus on the concepts to be included in the CONSORT and SPIRIT extensions for factorial trials.

3.4.2. Method

A two-day consensus meeting will be organised. Based on prior experience, we expect approximately 20-30 experts will attend the consensus meeting. The meeting will be held at the University of Nottingham, as it is accessible for many individuals and has several venues which are large enough to allow for appropriate social distancing if required. In the event of restrictions imposed by health pandemics or requirements of the panel members, the meeting will be held online via Microsoft Teams.

The consensus meeting will follow similar procedures to those in previous SPIRIT/CONSORT extensions[29, 33]. A brief background to the study will be presented along with results from the Delphi survey. Concepts that reached consensus for retention in the final guidelines (consensus in) and those that reached consensus for exclusions (consensus out) will be presented first to confirm their inclusion/exclusion. Concepts will be grouped in terms of their content (e.g., those relating to analysis, those relating to randomisation, etc) and presented to the group by a specialist in the area. Expert consensus group members will be encouraged to share their viewpoints relevant to these items and final agreement will be sought.

Discussion of concepts that failed to reach consensus (no consensus) in the Delphi survey will follow a similar process. Stakeholders will be encouraged to discuss the issues surrounding these items, specifically whether items (i) address elements unique to factorial trials and (ii) reflect information that should be included in a minimum reporting set of items.

3.4.3. Analysis

Key discussion points will be minuted whilst maintaining participant anonymity. No formal statistical analysis will take place. Consensus on the concepts for inclusion in the SPIRIT and CONSORT documents will be determined at the meeting through voting, with the criteria set at >50% agreement.

3.5. Phase Four: Knowledge Translation

3.5.1. Objective

Develop SPIRIT and CONSORT extensions based on the concepts agreed at the consensus meeting. Disseminate these documents, including guidance statements and an “Explanation and Elaboration” paper, which will provide model examples and an overview of design, conduct and analysis issues in factorial trials.

3.5.2. Activities

Once a final set of concepts that are relevant to factorial trials have been agreed at the consensus meeting, the study team will discuss and agree the wording of these items for the SPIRIT and CONSORT checklists, associated guidance statements, and the Explanation and Elaboration

document. These working documents will be developed alongside the main study papers (CONSORT and SPIRIT Statements and a shared Explanation and Elaboration document). The study team will agree the main outline of the papers prior to first drafts.

The working documents (checklists, guidance statements, and Explanation and Elaboration document) will be circulated to the wider group of participants from the consensus meeting for review prior to submission for publication. The Explanation and Elaboration document will include an explanation of each item, along with examples where appropriate. This document will act as a user-guide for the extensions and therefore should be comprehensible by individuals without specific expertise in factorial trials. For example, we will invite graduate students on trial-related courses to review the checklists and provide feedback on comprehension and usability. Further details on the dissemination plan are presented in the Dissemination and Implementation section.

4. Withdrawal Procedures

Participants may withdraw from further participation at any phase. Once data have been submitted via the DelphiManger software data will not be deleted.

5. Data Management Plan (DMP)

5.1. Data Description

Phase 1 of the study is a literature scoping review, which resulted in a bibliographic database of 29 full text articles. From these articles, 31 recommendations for consideration in the SPIRIT and/or CONSORT extensions were extracted in a tabulated form. These data are not further referred to in the DMP.

Phase 2 is an online survey. Participants will be required to provide personal information such as name, contact information and professional role. The survey items will ask participants to rate the perceived importance of a list of items relating to reporting of factorial trials. The approximate number of items that will be rated in round 1 is 50 (relating to the 31 concepts, 18 apply to SPIRIT and CONSORT, 13 apply to CONSORT only. Participants will be given the opportunity to provide free text responses, including suggesting additional items for rating in round 2 and round 3.

Phase 3 is a consensus meeting. Discussion points will not be attributed to individuals and no direct quotes will be reported. Therefore, this phase is not referred to further in this document.

Phase 4 is a knowledge translation phase. SPIRIT and CONSORT extensions will be developed alongside an Explanation and Elaboration document. These will be made freely available and are not referred to further in the DMP.

5.2. Data Collection and Management

Phase 2 survey data will be collected using DelphiManager software, a web-based system for building and managing Delphi surveys (http://www.cometinitiative.org/DelphiManager/index.html) at the University of Liverpool. DelphiManager requires mandatory registration for participants and collects name and email address to enable automated email reminders. It also contains a consent checkbox for compliance with GDPR. Survey creation and access to data by the survey owner is by unique username and password. Survey results are downloaded by the survey owner as CSV files, to enable analysis in statistical software such as Stata. Once downloaded identifying names will be cut from the data file and stored separately with a unique identifier. We anticipate there will be around 50-150 records, each corresponding to one survey participant.

5.3. Data Storage and Security

DelphiManager is a bespoke piece of software written in C# 4.0 using MVC 4. The data are stored in a MySQL database. The system itself is hosted on a dedicated server in the University of Liverpool data centre. Physical access to the University of Liverpool datacentre is restricted to authorised personnel. The DelphiManager server sits behind the University firewall and uses a 256 bit SSL certificate. Only authorised CTRC Information Systems (IS) team members have access to the server. All significant data stored within DelphiManager is encrypted. Access to the DelphiManager software instance is set up by the IS team and then password protected administrator accounts are managed by the study team. Any data entered into the individual instance of DelphiManager can be extracted by the study team administrators. CSV data files downloaded to UoN will be secured using an appropriate level of encryption and stored in a project-specific folder on a network drive, access to which will be restricted to necessary users and by user login/password. The drive will be backed up each night (incremental during the week and full over the weekend); this data store and backup will be provided by UoN Information Services.

5.4. Data Sharing and Access

Phase 2 data will be suitable for sharing as CSV files. The dataset(s) will have all personal information from participants removed and will contain only responses to the main survey questions, none of which contain any identifiable information. Summary information will be available on the study and Nottingham Clinical Trials Unit (NCTU) websites. Data sharing will be governed by NCTU's data sharing policy, which adopts a system of controlled access rather than open access to individual participant data. Data from this study will therefore not be deposited in an open repository but will be held by NCTU and shared upon approval of requests. Data sharing will not be unduly delayed; however, there will be a period of exclusivity of use by the research team of up to 12 months after the study end to enable all analyses and reporting to be completed.

5.5. Data Archiving

In compliance with the ICH/GCP guidelines, regulations and in accordance with the University of Nottingham Code of Research Conduct and Research Ethics, the Chief Investigator will maintain all records and documents regarding the conduct of the study. These will be retained for at least 7 years or for longer if required. If the responsible investigator is no longer able to maintain the study records, a second person will be nominated to take over this responsibility.

Study documents held by the Chief Investigator shall be finally archived at secure archive facilities at the University of Nottingham. This archive shall include all associated databases and any meta-data encryption codes. University of Nottingham data archive is underpinned by commercial digital storage which is audited on a twice-yearly basis for compliance with the ISO 27001 standard.

6. Statistical Considerations

Given the qualitative nature of the study, no formal statistical procedures will be employed.

7. Study Organisational Structure

7.1. Management Team

The study will be co-led by Alan Montgomery (AM: Professor of Medical Statistics and Clinical Trials and Director of Nottingham CTU) and Brennan Kahan (BK: Senior Research Fellow, MRC CTU at UCL), both of whom have extensive experience of applied and methodological research in the design, conduct, analysis, and reporting of factorial trials. AM and BK will provide line

supervision to the Research Fellow (Dr Sophie Hall, SH: University of Nottingham). Full team meetings will be held monthly, including all co-investigators where possible and progress reports shared. The meetings will be co-chaired by AM and BK with content and reports prepared by an administrator (Megan Birchenall, University of Nottingham) in collaboration with SH.

7.2. Trials Unit

The study is co-ordinated by the Nottingham Clinical Trials Unit.

7.3. Finance

The study is funded by the Medical Research Council (MRC). The funding commences on 1st September 2021. The costings include staff and investigators time, travel and sustenance for meetings and other consumables.

8. Ethical Considerations

The study will be conducted in compliance with the protocol, GCP and the applicable regulatory requirements. The study will be performed in accordance with the recommendations guiding physicians in biomedical research involving human participants, adopted by the 18th World Medical Association General Assembly, Helsinki, Finland, June 1964, amended at the 48th World Medical Association General Assembly, Somerset West, Republic of South Africa, October 1996 (website: http://www.wma.net/en/30publications/10policies/b3/index.html). The protocol will be submitted to and approved by the University of Nottingham Medical School Research Ethics Committee.

8.1. Ethical Issues

The study will recruit from a research audience. PPI members will be experienced in participating in academic research/clinical trials. The study will not involve vulnerable populations. All participants will be provided with a Participant Information Sheet, which will include information on the purpose of the data collected, data usage and sharing as well as withdrawal. Only anonymised data will be archived and shared. Participants will be informed that they can withdraw at any point from the study. Once data have been submitted it will not be possible to withdraw that information from the study. Data will be used for research purposes only. The research questions do not involve collecting data of a potentially sensitive nature.

8.2. Confidentiality and Data Protection

Personal data recorded on all documents will be regarded as strictly confidential and will be handled and stored in accordance with the Data Protection Act 2018. The Trials Office will maintain the confidentiality of all participant's data and will not disclose information by which participants may be identified to any third party.

9. Insurance and Indemnity

The University of Nottingham has appropriate and typical insurance coverage in place (including, but not limited to Clinical Trials, Professional Indemnity, Employer's Liability and Public Liability policies) in relation to the Institution's Legal Liabilities arising from the University's activities and those of its staff, whilst conducting University business and research activity.

10. Publication Policy

Results of this study will be submitted for publication in peer reviewed journals. The manuscript will be prepared by the study team authorship will be determined by mutual agreement. Authors must

acknowledge that the study was performed with the support of the MRC and NIHR CTU Support Funding.

11. Reference List

1. Chan, A.-W., et al., SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of Internal Medicine*, 2013. **158**(3): p. 200-207.
2. Schulz, K.F., D.G. Altman, and D. Moher, CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Trials*, 2010. **11**(1): p. 1-8.
3. Consort. Extensions of the Consort Statement. Available from: http://www.consortstatement.org/extensions.
4. Campbell, M.K., et al., Consort 2010 statement: extension to cluster randomised trials. *BMJ*, 2012. **345**.
5. Piaggio, G., et al., Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA*, 2012. **308**(24): p. 2594-2604.
6. Pandis, N., et al., CONSORT 2010 statement: extension checklist for reporting within person randomised trials. *British Journal of Dermatology*, 2019. **180**(3): p. 534-552.
7. Juszczak, E., et al., Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 statement. *JAMA*, 2019. **321**(16): p. 1610-1620.
8. Zwarenstein, M., et al., Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*, 2008. **337**.
9. Vohra, S., et al., CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *BMJ*, 2015. **350**.
10. Eldridge, S.M., et al., CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ*, 2016. **355**.
11. Dwan, K., et al., CONSORT 2010 statement: extension to randomised crossover trials. *BMJ*, 2019. **366**.
12. Porcino, A.J., et al., SPIRIT extension and elaboration for n-of-1 trials: SPENT 2019 checklist. *BMJ*, 2020. **368**.
13. Calvert, M., et al., Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA*, 2018. **319**(5): p. 483-494.
14. Kahan, B.C., et al., *Reporting of randomized factorial trials was frequently inadequate. Journal of Clinical Epidemiology*, 2020. **117**: p. 52-59.
15. Montgomery, A.A., M.P. Astin, and T.J. Peters, Reporting of factorial trials of complex interventions in community settings: a systematic review. *Trials*, 2011. **12**(1): p. 1-7.
16. Kasenda, B., et al., Prevalence, characteristics, and publication of discontinued randomized trials. *JAMA*, 2014. **311**(10): p. 1045-1052.
17. Briel, M., et al., Comparison of randomized controlled trials discontinued or revised for poor recruitment and completed trials with the same research question: a matched qualitative study. *Trials*, 2019. **20**(1): p. 1-12.
18. Walters, S.J., et al., Recruitment and retention of participants in randomised controlled trials: a review of trials funded and published by the United Kingdom Health Technology Assessment Programme. *BMJ Open*, 2017. **7**(3): p. e015276.
19. Williams, R.J., et al., Terminated trials in the ClinicalTrials. gov results database: evaluation of availability of primary outcome data and reasons for termination. *PloS one*, 2015. **10**(5): p. e0127242.
20. McAlister, F.A., et al., *Analysis and reporting of factorial trials: a systematic review. JAMA*, 2003. **289**(19): p. 2545-2553.
21. Montgomery, A.A., T.J. Peters, and P. Little, *Design, analysis and presentation of factorial randomised controlled trials. BMC Medical Research Methodology*, 2003. **3**(1): p. 1-5.

22. Kahan, B.C., Bias in randomised factorial trials. *Statistics in Medicine*, 2013. **32**(26): p. 4540-4549.
23. Freidlin, B. and E.L. Korn, Two-by-Two factorial cancer treatment trials: is sufficient attention being paid to possible interactions? *Journal of the National Cancer Institute*, 2017. **109**(9).
24. Clinical Trials. Available from: https://clinicaltrials.gov/.
25. Cro, S., et al., Evidence of unexplained discrepancies between planned and conducted statistical analyses: a review of randomised trials. *BMC Medicine*, 2020. **18**: p. 1-8.
26. Okoli, C. and S.D. Pawlowski, The Delphi method as a research tool: an example, design considerations and applications. *Information & Management*, 2004. **42**(1): p. 15-29.
27. Birko, S., E.S. Dove, and V. Özdemir, Evaluation of nine consensus indices in Delphi foresight research and their dependency on Delphi survey characteristics: a simulation study and debate on Delphi design and interpretation. *PloS one,* 2015. **10**(8): p. e0135162.
28. De Villiers, M.R., P.J. De Villiers, and A.P. Kent, The Delphi technique in health sciences education research. *Medical Teacher*, 2005. **27**(7): p. 639-643.
29. Imran, M., et al., Methods and results used in the development of a consensus-driven extension to the Consolidated Standards of Reporting Trials (CONSORT) statement for trials conducted using cohorts and routinely collected data (CONSORT-ROUTINE). *BMJ Open*, 2021. **11**(4): p. e049093.
30. Trevelyan, E.G. and N. Robinson, Delphi methodology in health research: how to do it? *European Journal of Integrative Medicine*, 2015. **7**(4): p. 423-428.
31. *DelphiManager*. Available from: https://www.comet-initiative.org/delphimanager/
32. Kwakkenbos, L., et al., Protocol for the development of a CONSORT extension for RCTs using cohorts and routinely collected health data. *Research Integrity and Peer Review*, 2018. **3**(1): p. 1-9.

12. Appendix 1

Table A1. Papers in the Scoping Review

1	Allore, H. G. & T. E. Murphy (2008). An examination of effect estimation in factorial and standardly-tailored designs. *Clinical Trials 5*(2): 121-130.
2	Baker, T. B., et al. (2017). Implementing Clinical Research Using Factorial Designs: A Primer. *Behavioural Therapy 48*(4): 567-580.
3	Bria, E., et al. (2006). Factorial design for randomized clinical trials. *Annals of Oncology 17*(10): 1607-1608.
4	Brittain, E. & J. Wittes (1989). Factorial designs in clinical trials: the effects of non-compliance and subadditivity. *Statistics in Medicine 8*(2): 161-171.
5	Byar, D. P. (1989). Some statistical considerations for design of cancer prevention trials. *Preventive Medicine 18*(5): 688-699.
6	Byar, D. P., et al. (1993). Incomplete factorial designs for randomized clinical trials. *Statistics in Medicine 12*(17): 1629-1641.
7	Byth, K. & V. Gebski (2004). Factorial designs: a graphical aid for choosing study designs accounting for interaction. *Clinical Trials 1*(3): 315-325.
8	Cairns, J., et al. (1991). Issues in the early termination of the aspirin component of the Physicians' Health Study. Data Monitoring Board of the Physicians' Health Study. *Annals of Epidemiology 1*(5): 395-405.
9	Collins, L. M., et al. (2014). Factorial experiments: efficient tools for evaluation of intervention components. *American Journal of Preventive Medicine 47*(4): 498-504.
10	Crespi, C. M. (2016). Improved Designs for Cluster Randomized Trials. *Annual Review of Public Health 37*: 1-16.
11	Curran, D., et al. (1999). Sample size estimation in phase III cancer clinical trials. *European Journal of Surgical Oncology 25*(3): 244-250.
12	Dakin, H. & A. Gray (2017). Economic evaluation of factorial randomised controlled trials: challenges, methods and recommendations. *Statistics in Medicine 36*(18): 2814-2830.
13	Dakin, H. A., et al. (2018). Partial factorial trials: comparing methods for statistical analysis and economic evaluation. *Trials 19*(1): 442.
14	Foley, R. N. (2009). Analysis of randomized controlled clinical trials. *Methods in Molecular Biology 473*: 113-126.
15	Freidlin, B. & E. L. Korn (2017). Two-by-Two Factorial Cancer Treatment Trials: Is Sufficient Attention Being Paid to Possible Interactions? *Journal of the National Cancer Institute 109*(9).
16	Green, S., et al. (2002). Factorial design considerations. *Journal of Clinical Oncology 20*(16): 3424-3430.
17	Green, S. B. (2000). Hypothesis testing in clinical trials. *Hematology Oncology Clinics of North America 14*(4): 785-795, vii-viii.
18	Kahan, B. C. (2013). Bias in randomised factorial trials. *Statistics in Medicine 32*(26): 4540-4549.
19	Korn, E. L. & B. Freidlin (2016). Non-factorial analyses of two-by-two factorial trial designs. *Clinical Trials 13*: 651-659.
20	Larntz, K., et al. (1996). Data analysis issues for protocols with overlapping enrollment. *Statistics in Medicine 15*(21-22): 2445-2453.

21	Lubsen, J. & S. J. Pocock (1994). Factorial trials in cardiology: pros and cons. *European Heart Journal 15*(5): 585-588.
22	McAlister, F. A., et al. (2003). Analysis and reporting of factorial trials: a systematic review. *JAMA 289*(19): 2545-2553.
23	McClure, L. A., et al. (2013). Monitoring futility in a two-by-two factorial design: the SPS3 experience. *Clinical Trials 10*(2): 250-256.
24	Mdege, N. D., et al. (2014). The 2 x 2 cluster randomized controlled factorial trial design is mainly used for efficiency and to explore intervention interactions: a systematic review. *Journal of Clinical Epidemiology 67*(10): 1083-1092.
25	Montgomery, A. A., et al. (2011). Reporting of factorial trials of complex interventions in community settings: a systematic review. *Trials 12*: 179.
26	Montgomery, A. A., et al. (2003). Design, analysis and presentation of factorial randomised controlled trials. *BMC Medical Research Methodology 3*: 26.
27	Pocock, S. J., et al. (2015). Challenging Issues in Clinical Trial Design: Part 4 of a 4-Part Series on Statistics for Clinical Trials. *Journal of the American College of Cardiology 66*(25): 2886-2898.
28	Kahan, B. C. et al. (2020). Reporting of randomized factorial trials was frequently inadequate. *Journal of Clinical Epidemiology 117*: 52-59.
29	Juszczak, E., et al. (2019). Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 statement. *JAMA 321*(16): 1610-1620.

13. Appendix 2

Table A2. Reporting Recommendations Forming the Delphi Items

Name	HelpText	DomainName
SPIRIT: Identification as a randomised factorial trial in the title	Including 'factorial' in the title helps to facilitate rapid judgements of relevance during literature searches.	Title
CONSORT: Identification as a randomised factorial trial in the title	Including 'factorial' in the title helps to facilitate rapid judgements of relevance during literature searches.	Title
SPIRIT: Scientific background and rationale for using a factorial design		Background and Rationale
CONSORT: Scientific background and rationale for using a factorial design		Background and Rationale
SPIRIT: Justification for whether an interaction is expected or not	Interactions can bias study conclusions, and understanding why interactions are/are not expected facilitates critical appraisal around the validity of the study design and methods	Background and Rationale
CONSORT: Justification for whether an interaction was expected or not	Interactions can bias study conclusions, and understanding why interactions are/are not expected facilitates critical appraisal around the validity of the study design and methods	Background and Rationale
SPIRIT: Specification of the research question(s) relating to the factorial design	Factorial designs allow investigators to address multiple research questions, including the effects of individual treatments, or whether treatments interact	Objectives

CONSORT: Specification of the research question(s) relating to the factorial design	Factorial designs allow investigators to address multiple research questions, including the effects of individual treatments, or whether treatments interact	Objectives
SPIRIT: Clear statement of the primary comparisons involved	Different comparisons can be made in factorial trials (e.g., all A vs. all not A; A alone vs. control; or A+B vs. double-control).	Objectives
CONSORT: Clear statement of the primary comparisons involved	Different comparisons can be made in factorial trials (e.g., all A vs. all not A; A alone vs. control; or A+B vs. double-control).	Objectives
SPIRIT: Type of factorial design (such as a full or partial factorial)	Different types of factorial designs require different study methods.	Trial Design
CONSORT: Type of factorial design (such as a full or partial factorial)	Different types of factorial designs require different study methods.	Trial Design
SPIRIT: Number of factors		Trial Design
CONSORT: Number of factors		Trial Design
SPIRIT: Number of levels within each factor		Trial Design
CONSORT: Number of levels within each factor		Trial Design

SPIRIT: The eligibility criteria for each factor, with any differences between the factors if applicable		Participants, Eligibility and Sample Size
CONSORT: The eligibility criteria for each factor, with any differences between the factors if applicable		Participants, Eligibility and Sample Size
SPIRIT: Planned sample size with details of how it was determined for each primary comparison	Different comparisons may require different sample sizes. A description of the sample size calculation for each comparison enables assessment of whether the study is adequately powered for each of its main objectives	Participants, Eligibility and Sample Size
CONSORT: Planned sample size with details of how it was determined for each primary comparison	Different comparisons may require different sample sizes. A description of the sample size calculation for each comparison enables assessment of whether the study is adequately powered for each of its main objectives	Participants, Eligibility and Sample Size
SPIRIT: Whether an interaction was assumed in the sample size calculation	Interactions affect the required sample size	Participants, Eligibility and Sample Size
CONSORT: Whether an interaction was assumed in the sample size calculation	Interactions affect the required sample size	Participants, Eligibility and Sample Size
SPIRIT: Stopping guidelines for each factor, with any differences between the factors if appropriate		Interventions, Monitoring and Stopping Guidelines
CONSORT: Stopping guidelines for each factor, with any differences between the factors if appropriate		Interventions, Monitoring and Stopping Guidelines

SPIRIT: Time-point of randomisation for each factor	Participants are sometimes allocated to different factors at different time-points, which may inform key study features	Sequence Generation
CONSORT: Time-point of randomisation for each factor	Participants are sometimes allocated to different factors at different time-points, which may inform key study features	Sequence Generation
SPIRIT: Description of estimand(s) for each primary and secondary outcome (treatment comparison, population, outcome definition, population-level summary, handling of intercurrent events)	Factorial designs allow estimation of different treatment effects (e.g., effect of A in the presence or absence of B). Specifying the estimand clarifies the target treatment effect	Statistical Methods and Analysis
CONSORT: Description of estimand(s) for each primary and secondary outcome (treatment comparison, population, outcome definition, population-level summary, handling of intercurrent events)	Factorial designs allow estimation of different treatment effects (e.g., effect of A in the presence or absence of B). Specifying the estimand clarifies the target treatment effect	Statistical Methods and Analysis
SPIRIT: Primary approach to statistical analysis (such as factorial, multi-arm) used to compare groups for primary and secondary outcomes, and details on how this approach will be chosen	Different analysis approaches may be taken (e.g., factorial vs. multi-arm), and the manner in which the approach is chosen (e.g., pre-specified, chosen based on trial data) can have implications for the risk of bias	Statistical Methods and Analysis
CONSORT: Primary approach to statistical analysis (such as factorial, multi-arm) used to compare groups for primary and secondary outcomes, and details on how this approach was chosen	Different analysis approaches may be taken (e.g., factorial vs. multi-arm), and the manner in which the approach is chosen (e.g., pre-specified, chosen based on trial data) can have implications for the risk of bias	Statistical Methods and Analysis
SPIRIT: How the other factor(s) will be handled during analysis	Adjustment for other factor(s) in a statistical model is sometimes desirable to improve efficiency or validity of the statistical analysis	Statistical Methods and Analysis

CONSORT: How the other factor(s) was/were handled during analysis	Adjustment for other factor(s) in a statistical model is sometimes desirable to improve efficiency or validity of the statistical analysis	Statistical Methods and Analysis
SPIRIT: Whether any adjustments for multiplicity will be applied and method used		Statistical Methods and Analysis
CONSORT: Whether any adjustments for multiplicity were applied and method used		Statistical Methods and Analysis
SPIRIT: Method(s) used to evaluate evidence of statistical interactions	Interactions can bias study conclusions and appropriate evaluation of such interactions is essential to study interpretation	Statistical Methods and Analysis
CONSORT: Method(s) used to evaluate evidence of statistical interactions	Interactions can bias study conclusions and appropriate evaluation of such interactions is essential to study interpretation	Statistical Methods and Analysis
SPIRIT: Likely impact of potential interactions on interpretation	Interpretation of factorial trials may depend on whether an interaction is identified	Statistical Methods and Analysis
CONSORT: Likely impact of identified interactions on interpretation	Interpretation of factorial trials may depend on whether an interaction is identified	Statistical Methods and Analysis
CONSORT: For each primary comparison, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	In factorial trials it can be difficult for readers to identify the relevant study flow, as this information may differ across primary comparisons. Presenting this information for each primary comparison increases clarity.	Participant Flow and Recruitment

CONSORT: For each primary comparison, losses and exclusions after randomisation, together with reasons	In factorial trials it can be difficult for readers to identify the relevant study flow, as this information may differ across primary comparisons. Presenting this information for each primary comparison increases clarity.	Participant Flow and Recruitment
CONSORT: Dates defining the periods of recruitment and follow-up; if different across factors, describe reason(s) for the differences and any statistical implications	One factor may have a short-term outcome, another a long term outcome, therefore follow-up periods may differ. If recruitment to a factor is affected by e.g., drug supply or is terminated early, then again, this should be reported.	Participant Flow and Recruitment
CONSORT: A table showing baseline demographic and clinical characteristics for each primary comparison	In factorial trials it can be difficult for readers to identify the relevant baseline characteristics, as this information may differ across primary comparisons. Presenting this information for each primary comparison increases clarity	Baseline Data and Numbers Analysed
CONSORT: For each primary comparison, the number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	In factorial trials it can be difficult for readers to identify the relevant baseline characteristics, as this information may differ across primary comparisons. Presenting this information for each primary comparison increases clarity	Baseline Data and Numbers Analysed
CONSORT: For each primary and secondary outcome, results for each primary comparison, the estimated effect size and its precision (such as 95% confidence interval)		Outcomes, Estimation and Ancillary Analysis
CONSORT: For each primary and secondary outcome, the estimated interaction effect and its precision	Interactions can bias study conclusions and evaluation of such interactions is essential to study interpretation	Outcomes, Estimation and Ancillary Analysis
CONSORT: Outcome data (including primary and secondary outcomes, harms, and adherence) presented by multi-arm group	Presentation by multi-arm groups (A, B, A+B, and double-control) allows readers to assess to what extent treatment comparisons or descriptive statistics for such data may be influenced by interactions	Outcomes, Estimation and Ancillary Analysis

CONSORT: All important harms or unintended effects in each primary comparison		Harms and Limitations
CONSORT: Influence of potential interactions	Interactions can bias study conclusions and considering the influence of such interactions is essential to study interpretation	Harms and Limitations
CONSORT: Whether adherence to intervention might have been affected by inclusion of other factors	When multiple interventions are tested together, one intervention may affect adherence to the other	Harms and Limitations
CONSORT: Effect of multiplicity of analyses (if relevant)		Harms and Limitations