JAMA | Special Communication

Reporting of Factorial Randomized Trials
Extension of the CONSORT 2010 Statement

Brennan C. Kahan, PhD; Sophie S. Hall, PhD; Elaine M. Beller, MAppStat; Megan Birchenall, BSc; An-Wen Chan, MD, DPhil; Diana Elbourne, PhD; Paul Little, MD; John Fletcher, MPH; Robert M. Golub, MD; Beatriz Goulao, PhD; Sally Hopewell, DPhil; Nazrul Islam, PhD; Merrick Zwarenstein, MBBCh, PhD; Edmund Juszczak, MSc; Alan A. Montgomery, PhD

IMPORTANCE Transparent reporting of randomized trials is essential to facilitate critical appraisal and interpretation of results. Factorial trials, in which 2 or more interventions are assessed in the same set of participants, have unique methodological considerations. However, reporting of factorial trials is suboptimal.

OBJECTIVE To develop a consensus-based extension to the Consolidated Standards of Reporting Trials (CONSORT) 2010 Statement for factorial trials.

DESIGN Using the Enhancing the Quality and Transparency of Health Research (EQUATOR) methodological framework, the CONSORT extension for factorial trials was developed by (1) generating a list of reporting recommendations for factorial trials using a scoping review of methodological articles identified using a MEDLINE search (from inception to May 2019) and supplemented with relevant articles from the personal collections of the authors; (2) a 3-round Delphi survey between January and June 2022 to identify additional items and assess the importance of each item, completed by 104 panelists from 14 countries; and (3) a hybrid consensus meeting attended by 15 panelists to finalize the selection and wording of items for the checklist.

FINDINGS This CONSORT extension for factorial trials modifies 16 of the 37 items in the CONSORT 2010 checklist and adds 1 new item. The rationale for the importance of each item is provided. Key recommendations are (1) the reason for using a factorial design should be reported, including whether an interaction is hypothesized, (2) the treatment groups that form the main comparisons should be clearly identified, and (3) for each main comparison, the estimated interaction effect and its precision should be reported.

CONCLUSIONS AND RELEVANCE This extension of the CONSORT 2010 Statement provides guidance on the reporting of factorial randomized trials and should facilitate greater understanding of and transparency in their reporting.

JAMA. 2023;330(21):2106-2114. doi:10.1001/jama.2023.19793

Supplemental content

CME at jamacmelookup.com

Related article at jamanetworkopen.com

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Brennan C. Kahan, PhD, MRC Clinical Trials Unit at University College London, 90 High Holborn, London WC1V 6LJ, United Kingdom (b.kahan@ucl.ac.uk).

In a factorial trial, 2 or more interventions are assessed in a single study by randomizing participants to multiple factors.[1-14] In a 2 × 2 trial with factors A and B, participants are randomized to receive intervention A or its comparator and also to intervention B or its comparator, meaning participants are assigned to 1 of 4 treatment groups: A alone, B alone, A plus B, or neither A nor B (**Table 1**).

Factorial designs are used to address different research questions (ie, estimands; **Box 1**). They can be used to evaluate more than 1 intervention in a single trial without increasing the sample size ("2-in-1" trials), to evaluate whether interventions interact, or to identify the best combination of interventions.[8,13,15,16] These disparate aims require different methodology, including sample size calculations and analysis strategies. Factorial trials also have additional methodological complexities compared with other trial designs, including choice of what treatment groups to include in main comparisons, how potential interactions should be handled during analysis, and nonconcurrent enrollment of participants.[1,2,4,6,10-13,17]

An extension of the Consolidated Standards of Reporting Trials (CONSORT) 2010 checklist for the reporting of factorial trials is presented in this article.[18,19] A glossary of key terms is provided in Box 1.

Methods

This CONSORT extension development occurred in parallel with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) extension for factorial trials.[20] First, we performed a scoping review using a MEDLINE search from inception to May 2019 to create an initial list of reporting recommendations applicable to factorial trials. Second, we performed

a 3-round Delphi survey (January to June 2022; 104 panelists from 14 countries) to identify additional items and assess the importance of each item. Third, an expert consensus meeting (September 6-7, 2022; 15 panelists) was held to establish the final checklist. Item wording was finalized after the meeting through iterative discussions.

Results

The checklist for the reporting of factorial randomized trials includes 16 modified items and 1 new item (**Table 2**). Reporting items for abstracts of factorial randomized trials are provided in **Table 3**.[21,22]

The scoping review identified 31 recommendations pertinent to reporting factorial trials, which were evaluated in the Delphi survey. Thirty-two recommendations met the criteria to be evaluated at the consensus meeting (1 recommendation was added in round 2 of the Delphi survey).

Given the variation in terminology used to describe factorial trials, items in this statement have been written to replace the original CONSORT items. Users are advised to refer to definitions of key terms in Box 1. This article contains brief explanations of the modified items in the CONSORT factorial extension. Details for interpretation of each item, and examples of good reporting, will be presented in a separate "explanation and elaboration" article.

Table 1. Example of a 2 × 2 Factorial Randomized Trial

	Treatment B (high-dose)[a,b]	Treatment B (low-dose)[a,b]
Treatment A (active)[a,b]	Active A + high-dose B[c]	Active A + low-dose B[c]
Treatment A (placebo)[a,b]	Placebo A + high-dose B[c]	Placebo A + low-dose B[c]

[a] A and B are factors.

[b] Active A and placebo A are levels within factor A and high-dose B and low-dose B are levels within factor B. Low-dose B is taken as the control condition for factor B.

[c] Active A plus high-dose B, active A plus low-dose B, placebo A plus high-dose B, and placebo A plus low-dose B are the treatment groups. In a "full" factorial trial all participants are eligible to be randomized between each of the 4 treatment groups; in a "partial" factorial trial, a subset of participants would only be randomized between high- and low-dose B and assigned to placebo A without randomization. In a "factorial" analysis, all participants allocated to intervention A (active A plus low-dose B and active A plus high-dose B) are compared against those not allocated to A (placebo A plus low-dose B and placebo A plus high-dose B), and similarly for the comparison for intervention B. In a "multiarm" analysis, each of the treatment groups are compared against control (eg, active A plus high-dose B, active A plus low-dose B, and placebo A plus high-dose B are all compared against placebo A plus low-dose B).

CONSORT Checklist Extension for Factorial Randomized Trials

Item 1a. CONSORT 2010: Identification as a randomized trial in the title

Extension for factorial trials: Identification as a factorial randomized trial in the title | Notifying readers of the factorial design alerts them to potential implications of the design for analysis and interpretation.[2,4,5,8,23,24]

Item 2a. CONSORT 2010: Scientific background and explanation of rationale

Extension for factorial trials: Rationale for using a factorial design, including whether an interaction is hypothesized | Different research hypotheses require different methodology. By clarifying the rationale for using the factorial design, as well as whether an interaction is hypothesized, readers are signposted toward the key objectives and alerted to the assumptions and methodological features required.[1,4-6,24]

Item 2b. CONSORT 2010: Specific objectives or hypotheses

Extension for factorial trials: A statement of which treatment groups form the main comparisons | In factorial trials, interventions can be compared in different ways. In a 2 × 2 factorial trial with factors A and B, the treatment effect for intervention A vs its comparator can be estimated by comparing (1) participants randomized to receive A vs not A; (2) those randomized to receive A alone vs neither A nor B; or (3) those randomized to receive A plus B vs B alone. These alternative comparisons can target different estimands and are underpinned by different assumptions (**Box 2**).[4,6,11] An estimand describes the target treatment effect to be estimated from the trial.

Item 3a. CONSORT 2010: Description of trial design (such as parallel, factorial) including randomization ratio

Extension for factorial trials: Description of the type of factorial trial (such as a full or partial, number of factors, and levels within each factor) | Most factorial trials use a "full" factorial design, whereby all participants are eligible to be randomized to all combinations of factors and factor levels.[9,25,26] Other designs include "fractional" factorial designs (where some combinations of factors are omitted) and "partial" factorial designs (where some participants are only eligible to be randomized to certain factors), which require alternative methodology.[1,27]

Item 4a. CONSORT 2010: Eligibility criteria for participants

Extension for factorial trials: Eligibility criteria for each factor, noting any differences, if applicable | Differences in eligibility criteria across factors can have implications for the design and analysis and can increase the risk of bias if not handled properly. For instance, participants who are not eligible for randomization to a specific factor should not be included in the comparison for that factor, because their inclusion means the analysis is no longer based on a randomized comparison, which can lead to confounding bias.[1,27]

Item 7a. CONSORT 2010: How sample size was determined

Extension for factorial trials: How sample size was determined for each main comparison, including whether an interaction was assumed in the calculation | Sample size calculations for factorial designs are more complicated than in standard parallel-group designs. In some factorial trials, the planned main comparisons may require different sample sizes if they are expected to produce different effect sizes or if the choice of primary outcome varies for each

Box 1. Glossary of Terms Used in the Extension of the CONSORT 2010 Statement

Comparison: What treatment groups will be compared against each other. For example, the effect of intervention A may be estimated by comparing all participants randomized to active A (treatment groups active A plus high-dose B and active A plus low-dose B) with all participants randomized to placebo A (treatment groups placebo A plus high-dose B and placebo A plus low-dose B).

Estimand: A description of the treatment effect to be estimated from the trial, including specification of the treatment conditions, population, end point, summary measure, and strategies to handle intercurrent events. Factorial trials should additionally specify how the other factor(s) are to be handled in the estimand (eg, whether interest lies in the effect of active A plus low-dose B vs placebo A plus low-dose B or else active A plus high-dose B vs placebo A plus high-dose B).

Factor: Each intervention and its comparator(s) together comprise a factor (eg, active A and placebo A together comprise one factor and high-dose B and low-dose B together make up the other factor).

Factorial analysis: Also called an "at-the-margins" analysis. All participants randomized to active A (treatment groups active A plus high-dose B and active A plus low-dose B) are compared against all those randomized to placebo A (placebo A plus high-dose B and placebo A plus low-dose B) and similarly for the factor B comparison.

Factorial trial: When 2 or more interventions are assessed in the same participants within a single study.

Fractional factorial design: Some combinations of factors are omitted. For example, in a trial with 3 factors (A, B, and C), participants may be randomized to 4 of the 8 possible combinations.

Full factorial design: All factors and levels are combined so the design comprises all possible combinations of factor levels and all participants are eligible to be randomized for each factor.

Interaction: Interactions occur when the effect of one treatment depends on whether participants also receive the other treatment (eg, active A may be less effective when used alongside high-dose B than when used with low-dose B). Interactions may occur for biological or social reasons (eg, if receipt of one treatment affects the mechanism of action for the other). Interactions may also occur due to choice of analysis scale (eg, active A may be equally effective with high-dose B as with low-dose B when measured on the risk ratio scale, but less effective on the risk difference scale). Trials interested in evaluating whether treatments interact are typically interested in biological/social interactions, while trials that use analyses that require an assumption of no interaction are affected by any type of interaction.

Level within factors: The specific interventions within a factor are the levels (eg, active A and placebo A are the 2 levels of factor A).

Main comparison(s): The comparison(s) that will primarily be used to draw conclusions about effectiveness of each intervention.

Multiarm analysis: Also called an "inside-the-table" analysis. The treatment groups active A plus low-dose B, placebo A plus high-dose B, and active A plus high-dose B are each compared against placebo A plus low-dose B (double-control).

Treatment group: The unique combinations of factors and levels to which participants can be randomized (eg, active A plus high-dose B comprise one treatment group and active A plus low-dose B another).

Partial factorial design: Some participants are not randomized to certain factors. For example, a subset of participants will only be randomized between active A vs control A and will receive control B automatically.

factor.[6,28] If an interaction is hypothesized, the sample size may need to be increased.[1,2,6,24]

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

Extension for factorial trials: When applicable, explanation of any interim analyses and stopping guidelines, noting any differences across main comparisons and reasons for differences | The plan for interim analyses and subsequent stopping guidelines may be different for each factor.[27] If one factor is stopped before the other, there may be implications for randomization, choice of comparator, or analysis.[1,27,29]

Item 8b. CONSORT 2010: Type of randomization; details of any restriction (such as blocking and block size)

Extension for factorial trials: If applicable, whether participants were randomized to factors at different points | Participants may be randomized to factors at different time points (eg, for factor A at diagnosis of disease then for factor B after treatment A is complete). The time point of randomization for each factor may inform key design features, such as the baseline period, duration of follow-up, and likelihood of treatments interacting.[2]

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

Extension for factorial trials: Statistical methods used for each main comparison for primary and secondary outcomes, including:

- **Whether the target treatment effect for each main comparison pertains to the effect in the presence or absence of other factors**
  The statistical methods alone are not always sufficient to allow readers to understand the exact treatment effect (estimand) being estimated.[30-32] In factorial trials, the treatment groups used for comparison are not always the same as those in which there is interest in estimating the treatment effect.[11,33] For example, many factorial trials use a factorial analysis to compare "all A" vs "all not A" for reasons of efficiency, even though interest really lies in the effect of A alone vs control (the effect of A in the absence of B) or, alternatively, the effect of A plus B vs B alone (the effect of A in the presence of B) if treatment B has been demonstrated to be effective.[11] A clear description of the target treatment effect, including whether it pertains to the effect in the presence or absence of other factors, allows readers to understand the exact question being addressed.[11,30,31,34]
- **Approach to analysis, such as factorial or multiarm**
  Different statistical methods can be used to analyze a factorial trial depending on the estimand of interest. In a factorial (or "at-the-margins") analysis, all participants randomized to factor

Table 2. CONSORT Checklist of Information to Include When Reporting Factorial Randomized Trials[a,b]

Section	Item No.	CONSORT 2010 Statement checklist item	Extension for factorial trials
Title and abstract			
Title	1a	Identification as a randomized trial in the title	Identification as a factorial randomized trial in the title
Abstract	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	See separate factorial checklist for abstracts
Introduction			
Background	2a	Scientific background and explanation of rationale	Scientific background and rationale for using a factorial design, including whether an interaction is hypothesized
Objectives	2b	Specific objectives or hypotheses	Specific objectives or hypotheses and a statement of which treatment groups form the main comparisons[b]
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Description of the type of factorial trial (such as full or partial, number of factors, levels within each factor[b]) and allocation ratio
Change from protocol	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	Eligibility criteria for each factor, noting any differences, if applicable
Setting and location	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	
Changes to outcomes	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size	7a	How sample size was determined	How sample size was determined for each main comparison, including whether an interaction was assumed in the calculation
Interim analyses and stopping guidelines	7b	When applicable, explanation of any interim analyses and stopping guidelines	When applicable, explanation of any interim analyses and stopping guidelines, noting any differences across main comparisons and reasons for differences
Randomization			
Sequence generation	8a	Method used to generate the random allocation sequence	
Sequence generation	8b	Type of randomization; details of any restriction (such as blocking and block size)	Type of randomization, details of any restriction (such as blocking and block size), and, if applicable, whether participants were randomized to factors at different time points
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes)	
Similarity of interventions	11b	If relevant, description of the similarity of interventions	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	Statistical methods used for each main comparison for primary and secondary outcomes, including: Whether the target treatment effect for each main comparison pertains to the effect in the presence or absence of other factors Approach to analysis, such as factorial or multiarm How the approach was chosen, such as prespecified or based on estimated interaction If factorial approach was used, whether factors were adjusted for each other If applicable, how nonconcurrent recruitment to factors was handled Method(s) used to evaluate statistical interaction(s)
Additional analyses	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	

(continued)

Table 2. CONSORT Checklist of Information to Include When Reporting Factorial Randomized Trials[a,b] (continued)

Section	Item No.	CONSORT 2010 Statement checklist item	Extension for factorial trials
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome	For each main comparison, the number of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome
Losses and exclusions	13b	For each group, losses and exclusions after randomization, together with reasons	For each main comparison, losses and exclusions after randomization, together with reasons
Recruitment	14a	Dates defining the periods of recruitment and follow-up	Dates defining the periods of recruitment and follow-up for each factor, noting any differences, with reasons
Trial end	14b	Why the trial ended or was stopped	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	A table showing baseline demographic and clinical characteristics for each main comparison
Numbers analyzed	16	For each group, the number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each main comparison, the number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group and the estimated effect size and its precision (such as 95% CI)	For each primary and secondary outcome, results for each main comparison, the estimated effect size, and its precision (such as 95% CI) For each primary outcome, the estimated interaction effect and its precision If done, the estimated interaction effects and precision for secondary outcomes
Binary outcomes	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
Ancillary analyses	18a	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory	
Additional data summaries[c]	18b		Participant flow, losses and exclusions, baseline data, and outcome data (including primary and secondary outcomes, harms, and adherence) presented by treatment groups[b]
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	All important harms or unintended effects for each main comparison
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	
Generalizability	21	Generalizability (external validity, applicability) of the trial findings	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
Other information			
Registration	23	Registration number and name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	

[a] It is strongly recommended that this checklist is read in conjunction with the CONSORT 2010 checklist (https://www.equator-network.org/reporting-guidelines/consort/) and Statement Explanation and Elaboration paper[18] for important clarification on the items. The CONSORT-factorial Checklist is licensed by the CONSORT-factorial Group under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

[b] Each overall intervention group to be compared is a *factor* (eg, in a 2 × 2 trial with factors A and B, active A and control A together comprise one factor and active B and control B together comprise another factor). The specific interventions within a factor are the *levels* (eg, active A and control A are the 2 levels of factor A). The unique combinations of factors and levels are *treatment groups* (eg, in a 2 × 2 trial with factors A and B there will be 4 treatment groups: active A plus control B, active A plus active B, etc). What treatment groups will be compared against each other to draw main conclusions about the effectiveness of each intervention is the *main comparison*.

[c] New item.

A (A alone and A plus B) are compared with all those not randomized to A (B alone and double-control).[2,4,6,11,35,36] Alternatively, in a multiarm (or "inside-the-table") analysis, the trial is analyzed as if a multiarm design had been used.[2,4-6,10-12,17,23,35,36] The 2 approaches offer different benefits and require different assumptions (Box 2).

- **How the approach was chosen, such as prespecified or based on estimated interaction**

Using a test of interaction to guide the choice of analysis can introduce bias and is not recommended.[17] Clarification of whether the final analysis approach was prespecified based on prior knowledge or an assumption of no interaction or chosen based on the

Table 3. Items to Include When Reporting a Randomized Factorial Trial in a Journal or Conference Abstract[a]

Item	CONSORT for abstracts checklist item	Extension for factorial trials
Title	Identification of the study as randomized	Identification of the study as a factorial randomized trial
Authors[b]	Contact details for the corresponding author	
Trial design	Description of the trial design (eg, parallel, cluster, noninferiority)	Description of the trial design (eg, parallel, cluster, noninferiority) and number of factors (eg, 2 × 2)
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	Eligibility criteria for each factor, noting any differences if applicable, and the settings where the data were collected
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	
Outcome	Clearly defined primary outcome for this report	
Randomization	How participants were randomized to interventions	
Blinding (masking)	Whether or not participants, caregivers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants randomized to each group	Number of participants randomized for each main comparison
Recruitment	Trial status	
Numbers analyzed	Number of participants analyzed in each group	Number of participants analyzed for each main comparison
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	For the primary outcome, results for each main comparison, the estimated effect size and its precision, and estimated interaction effect and its precision
Harms	Important adverse events	Important adverse events for each main comparison
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register	
Funding	Source of funding	

[a] The CONSORT-factorial Abstract Checklist is licensed by the CONSORT-factorial Group under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

[b] This item is specific to conference abstracts.

size of the estimated interaction helps alert readers to any risk of bias associated with the analysis approach.

- **Method(s) used to evaluate statistical interaction(s)**

It is recommended practice to evaluate the presence of statistical interactions, either because analyses rely on the assumption that treatments do not interact or because the interaction is itself of direct interest.[2,4-6,10,11,24] The presence of an interaction may depend on the scale of analysis (eg, an interaction may be present on the risk difference scale, but not the risk ratio scale), so careful consideration should be given to the choice of scale. Reporting details of how interaction(s) were evaluated, and on what scale, enables readers to understand the appropriateness of method(s).

- **If factorial approach used, whether factors were adjusted for each other**

Factorial analyses can be adjusted for whether participants were randomized to the other factor(s) by including a term for this in the statistical model.[2,6,11,28] This can increase statistical power and, in some cases, failure to adjust for the other factors can introduce bias for certain estimands.[11]

- **If applicable, how nonconcurrent recruitment to factors was handled**

Nonconcurrent recruitment, in which certain participants are not randomized for some factors (eg, if the trial used a partial factorial design or recruitment to one factor is paused or terminated), can induce bias if not handled correctly during analysis (see item 4a).[1,27]

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

Extension for factorial trials: For each main comparison, the number of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome | For factorial trials, especially those beyond a 2 × 2 design, it can be difficult for readers to identify the relevant participant flow because this information may differ across main comparisons. Presenting this information for each main comparison increases clarity and understanding.[2,4-6,8,10,35]

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

Extension for factorial trials: Dates defining the periods of recruitment and follow-up for each factor, noting any differences with reasons | If periods of recruitment are different across factors, participants enrolled after one factor has stopped recruitment

Box 2. Estimands for Factorial Trials

Estimands for factorial trials

An estimand describes a research question a trial sets out to address (Box 1).

Different types of estimands may be specified for factorial trials depending on the aims.

An estimand for the effect of treatment A could be defined based on a comparison of treatment A vs not A if no one received treatment B or as the effect of A vs not A if everyone received treatment B.

The former may be more common for "2-in-1" factorial trials because it provides the effect of treatment A that would be seen in a parallel-group design where treatment B is not used. However, either estimand may be of interest.

Alternatively, an estimand for treatment A could also be defined based on the effect of A vs not A averaged across those who do and those who do not receive treatment B.[a] Because this estimand does not typically reflect how treatments are used in practice, other choices are usually more relevant for 2-in-1 trials.

For trials in which the aim is to determine whether treatments interact, the estimand may be based around the difference between the effect of treatment A if no one received treatment B vs the effect if everyone received treatment B.

Implications for statistical analysis[b]

The method of statistical analysis should be determined by the estimand (ie, research question).

Two-in-1 trials typically use a factorial analysis because this realizes the efficiency gains inherent to the factorial design. However, because this analysis averages across the 2 strata of those randomized to receive and not receive B, it only estimates the "effect of treatment A if no one receives B" if treatments A and B do not interact. When treatments do interact, it estimates the mean effect of A across the strata of B. Therefore, assessment of the interaction is essential to determine whether the factorial analysis is estimating the desired estimand.

A multiarm ("inside-the-table") analysis could also be used to estimate the effect of treatment A if no one receives B, and is unbiased regardless of whether treatments A and B interact. However, it does not realize the efficiency gained through using a factorial design, so it is less frequently used for 2-in-1 trials.

[a] This averaging could correspond to the study proportions randomized to receive treatment B and not B or to some other proportions defined by the investigators. The exact method of determining the mean therefore needs to be made explicit.

[b] A factorial analysis can be used to estimate either (1) the effect of A if no one received B; (2) the effect of A if everyone received B; or (3) the effect of A averaged over those who received and did not receive B according to the study proportions. The first 2 of these estimates require the assumption of no interaction, but the analysis for the third does not. A multiarm analysis can be used to estimate by either comparing A alone vs double-control (as described above) or comparing A plus B vs B alone. These do not require the assumption of no interaction. If interest lies in the effect of A averaged over those who do and do not receive B according to proportions other than the study proportions, this could be estimated by first estimating the effect of A separately in both stratum (those who receive and do not receive B) then taking a weighted average of these according to the desired proportions. This analysis does not require the assumption of no interaction.[11]

will only be eligible to be randomized for the ongoing factor(s), posing similar statistical issues as in a partial factorial design (see CONSORT item 4a).[27]

Item 17a. CONSORT 2010: For each primary and secondary outcome, results for each group and the estimated effect size and its precision (such as 95% CI)

Extension for factorial trials: For each primary and secondary outcome, results for each main comparison, the estimated effect size, and its precision (such as 95% CI)

For each primary outcome, the estimated interaction effect and its precision

If done, estimated interaction effects and precision for secondary outcomes | For factorial trials predicated on the assumption of no interaction (2-in-1 trials) or those in which the interaction is of main interest, evaluation of interactions is essential to interpretation.[2,4-6,10,11,24] The size of the estimated interaction effect should be presented along with a measure of precision, such as the 95% CI.[2,5,6] For trials in which evaluation of interaction(s) is not deemed essential, this decision should be justified.

Item 18b. CONSORT 2010: New item (additional data summaries)

New item for factorial trials: Participant flow, losses and exclusions, and outcome data (including primary and secondary outcomes, harms, and adherence) presented by treatment groups | Outcomes and other postrandomization data such as adherence, harms, and participant flow may be affected when treatments interact.[26] Presentation of such data by treatment group (eg, groups A alone, B alone, A plus B, and double-control in a 2 × 2 trial), in addition to presentation by main comparisons, allows readers to assess to what extent such data may be unduly influenced by interactions due to the factorial design.[3-6,8,10]

Discussion

This extension to the CONSORT 2010 Statement provides guidance for reporting factorial trials. The extension checklist represents the minimum essential requirements for reporting of factorial trials; for some trials there will be additional items that are important to report. For instance, if primary or secondary outcomes differ by factor, this should be reported. Similarly, if multiple testing is deemed to be an issue, authors should report how this was handled.

This extension was developed in conjunction with the SPIRIT extension for factorial trials. Together, these guidelines provide a framework for cohesive reporting from the trial protocol to publication of results. The latest version of this and other CONSORT statements can be found online (https://www.equator-network.org/).

Limitations

This study has several limitations. First, this extension was developed for studies in which results for each factor would be published simultaneously in the same article. This may not always be feasible, for instance, due to the early stopping of one factor or because each factor requires different durations of follow-up. In this case, we recommend that each publication follows the checklist as far as possible, while recognizing that the information for some items might

differ. For example, each article could report how the sample size was determined for the relevant comparison, rather than the sample size calculations for each comparison (although each calculation would need to clarify whether an interaction was assumed).

Second, although the EQUATOR guidelines were followed to develop this guideline, Delphi respondents were self-selecting and consensus meeting panelists were purposively identified based on their expertise. Therefore, although results represent the views of a large, multinational group of experts and end users, the views of individuals not well represented by the Delphi survey or consensus meeting panelists may differ. However, the systematic and evidence-based approach used to develop this guideline, including a rigorous scoping review, should help mitigate the potential effects of these limitations.

Conclusion

This extension of the CONSORT 2010 Statement provides specific guidance for the reporting of factorial randomized trials to facilitate greater transparency and completeness in the reporting of these trials.

ARTICLE INFORMATION

Accepted for Publication: September 13, 2023.

Author Affiliations: MRC Clinical Trials Unit at UCL, London, United Kingdom (Kahan); Nottingham Clinical Trials Unit, School of Medicine, University of Nottingham, Nottingham, United Kingdom (Hall, Birchenall, Juszczak, Montgomery); Institute for Evidence-Based Healthcare, Bond University, Queensland, Australia (Beller); Women's College Research Institute, University of Toronto, Toronto, Ontario, Canada (Chan); London School of Hygiene and Tropical Medicine, London, United Kingdom (Elbourne); Primary Care Research Centre, School of Primary Care, Population Sciences and Medical Education, Faculty of Medicine, University of Southampton, Southampton, United Kingdom (Little, Islam); The BMJ, BMA House, Tavistock Square, London, United Kingdom (Fletcher, Islam); Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois (Golub); Health Services Research Unit, University of Aberdeen, Aberdeen, Scotland (Goulao); Oxford Clinical Trials Research Unit, University of Oxford, Oxford, United Kingdom (Hopewell); Centre For Studies in Family Medicine, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada (Zwarenstein).

Author Contributions: Drs Kahan and Hall had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Drs Kahan and Hall contributed equally and Mr Juszczak and Dr Montgomery contributed equally.
Concept and design: Kahan, Beller, Chan, Elbourne, Little, Fletcher, Golub, Hopewell, Juszczak, Montgomery.
Acquisition, analysis, or interpretation of data: Kahan, Hall, Beller, Birchenall, Chan, Elbourne, Fletcher, Golub, Goulao, Islam, Zwarenstein, Juszczak, Montgomery.
Drafting of the manuscript: Kahan, Hall, Elbourne, Little, Golub, Juszczak, Montgomery.
Critical review of the manuscript for important intellectual content: Hall, Beller, Birchenall, Chan, Elbourne, Little, Fletcher, Golub, Goulao, Hopewell, Islam, Zwarenstein, Juszczak, Montgomery.
Statistical analysis: Kahan, Hall, Beller.
Obtained funding: Kahan, Elbourne, Little, Juszczak, Montgomery.
Administrative, technical, or material support: Hall, Beller, Birchenall, Goulao.
Supervision: Kahan, Hall, Little, Montgomery.
Other - provided advice on method based on experience of previous CONSORT documents: Zwarenstein.

Conflict of Interest Disclosures: Dr Islam reported receiving remuneration for work as a research editor for *The BMJ* and receiving funding from the UK Office for National Statistics and UK National Institute for Health and Care Research. Dr Hopewell is a member of the SPIRIT-CONSORT executive group and leading the current update of the SPIRIT 2013 and CONSORT 2010 reporting guidelines funded by the UK Medical Research Council National Institute for Health Research Better Methods, Better Research (MR/W020483/1). Dr Fletcher reported receiving remuneration for work as an associate editor for *The BMJ*. Dr Montgomery reported receiving grants from National Institute of Health Research outside the submitted work. No other disclosures were reported.

Funding/Support: This work was supported by the Medical Research Council (grant No. MR/V020803/1).

Role of the Funder/Sponsor: The Medical Research Council had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: Dr Golub held the position of Executive Deputy Editor of *JAMA* during guideline development, but he was not involved in any of the decisions regarding review of the manuscript or its acceptance. This article reflects the views of the authors, the Delphi panelists, and the consensus meeting panelists and may not represent the views of the broader stakeholder groups, the authors' institutions, or other affiliations.

Data Sharing Statement: See the Supplement.

Additional Contributions: We thank and acknowledge the contributions of all members of the Delphi study, who were not compensated for their contributions, including the following Delphi survey contributors: Aaron Orkin, Aiping Lyu, Angela Fidler Pfammatter, Ben Cromarty, Catherine Hewitt, Christine Bond, Christopher Partlett, Christopher Schmid, Claire L. Chan, David Moher, Derrick Bennett, Elizabeth George, Evan Mayo-Wilson, Giovannino Ciccone, Graeme S. MacLennan, Halvor Sommerfelt, Hams Hamed, Helen Dakin, Himanshu Popat, Ian White, Jay Park, Jennifer Nicholas, Jonathan Emberson, Joseph C. Cappelleri, Julia Edwards, Julien Vos, Kath Starr, Kerry Dwan, Lee Middleton, Lehana Thabane, Lori Frank, Madelon van Wely, Marie-Joe Nemnom, Mark Hull, Martha Alejandra Morales-Sánchez, Martin Law, Martyn Lewis, Michael Forstner, Mike Bradburn, Monica Taljaard, Munya Dimairo, Nick Freemantle, Nuria Porta, Nurulamin Noor, Olalekan Lee Aiyegbusi, Patricia Logullo, Philip Pallmann, Ranjit Lall, Reuben Ogollah, Richard Haynes, Richard L. Kravitz, Robert Platt, Sarah Pirrie, Sharon Love, Shaun Treweek, Siobhan Creanor, Sunita Vohra, Susan Dutton, Suzie Cro, Tianjing Li, Tim Morris, Timothy Collier, Trish Hepburn, Vivian A. Welch, William Tarnow-Mordi, and Yolanda Barbachano.

REFERENCES

1. Green S, Liu PY, O'Sullivan J. Factorial design considerations. *J Clin Oncol*. 2002;20(16):3424-3430. doi:10.1200/JCO.2002.03.003

2. Kahan BC, Tsui M, Jairath V, et al. Reporting of randomized factorial trials was frequently inadequate. *J Clin Epidemiol*. 2020;117(117):52-59. doi:10.1016/j.jclinepi.2019.09.018

3. Lubsen J, Pocock SJ. Factorial trials in cardiology: pros and cons. *Eur Heart J*. 1994;15(5):585-588. doi:10.1093/oxfordjournals.eurheartj.a060552

4. McAlister FA, Straus SE, Sackett DL, Altman DG. Analysis and reporting of factorial trials: a systematic review. *JAMA*. 2003;289(19):2545-2553. doi:10.1001/jama.289.19.2545

5. Montgomery AA, Astin MP, Peters TJ. Reporting of factorial trials of complex interventions in community settings: a systematic review. *Trials*. 2011;12:179. doi:10.1186/1745-6215-12-179

6. Montgomery AA, Peters TJ, Little P. Design, analysis and presentation of factorial randomised controlled trials. *BMC Med Res Methodol*. 2003;3:26. doi:10.1186/1471-2288-3-26

7. Byth K, Gebski V. Factorial designs: a graphical aid for choosing study designs accounting for interaction. *Clin Trials*. 2004;1(3):315-325. doi:10.1191/1740774504cn026oa

8. Dakin H, Gray A. Economic evaluation of factorial randomised controlled trials: challenges, methods and recommendations. *Stat Med*. 2017;36(18):2814-2830. doi:10.1002/sim.7322

9. Dakin HA, Gray AM, MacLennan GS, et al. Partial factorial trials: comparing methods for statistical analysis and economic evaluation. *Trials*. 2018;19(1):442. doi:10.1186/s13063-018-2818-x

10. Freidlin B, Korn EL. Two-by-two factorial cancer treatment trials: is sufficient attention being paid to possible interactions? *J Natl Cancer Inst*. 2017;109(9):djx146. doi:10.1093/jnci/djx146

11. Kahan BC, Morris TP, Goulao B, et al. Estimands for factorial trials. *Stat Med*. 2022;41:4299-4310. doi:10.1002/sim.9510

12. Korn EL, Freidlin B. Non-factorial analyses of two-by-two factorial trial designs. *Clin Trials*. 2016; 13(6):651-659. doi:10.1177/1740774516659472

13. McClure LA, Coffey CS, Howard G. Monitoring futility in a two-by-two factorial design: the SPS3 experience. *Clin Trials*. 2013;10(2):250-256. doi:10.1177/1740774512474374

14. Lin DY, Gong J, Gallo P, et al. Simultaneous inference on treatment effects in survival studies with factorial designs. *Biometrics*. 2016;72(4):1078-1085. doi:10.1111/biom.12507

15. Bria E, Di Maio M, Nistico C, et al. Factorial design for randomized clinical trials. *Ann Oncol*. 2006;17(10):1607-1608. doi:10.1093/annonc/mdl106

16. Byar DP. Some statistical considerations for design of cancer prevention trials. *Prev Med*. 1989;18(5):688-699. doi:10.1016/0091-7435(89)90040-6

17. Kahan BC. Bias in randomised factorial trials. *Stat Med*. 2013;32(26):4540-4549. doi:10.1002/sim.5869

18. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2012;10(1):28-55. doi:10.1136/bmj.c869

19. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332. doi:10.1136/bmj.c332

20. Hall SS, Juszczak E, Birchenall M, et al. Development of extensions to SPIRIT and CONSORT guidelines: the Reporting Factorial Trials (RAFT) study. OSF. Preprint posted online April 18, 2023. https://osf.io/undpm

21. Hopewell S, Clarke M, Moher D, et al; CONSORT Group. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet*. 2008;371(9609):281-283. doi:10.1016/S0140-6736(07)61835-2

22. Hopewell S, Clarke M, Moher D, et al; CONSORT Group. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med*. 2008;5(1):e20. doi:10.1371/journal.pmed.0050020

23. Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 Statement. *JAMA*. 2019;321(16):1610-1620. doi:10.1001/jama.2019.3087

24. Mdege ND, Brabyn S, Hewitt C, Richardson R, Torgerson DJ. The 2 × 2 cluster randomized controlled factorial trial design is mainly used for efficiency and to explore intervention interactions: a systematic review. *J Clin Epidemiol*. 2014;67(10):1083-1092. doi:10.1016/j.jclinepi.2014.06.004

25. Allore HG, Murphy TE. An examination of effect estimation in factorial and standardly-tailored designs. *Clin Trials*. 2008;5(2):121-130. doi:10.1177/1740774508089278

26. Baker TB, Smith SS, Bolt DM, et al. Implementing clinical research using factorial designs: a primer. *Behav Ther*. 2017;48(4):567-580. doi:10.1016/j.beth.2016.12.005

27. White IR, Choodari-Oskooei B, Sydes MR, et al. Combining factorial and multi-arm multi-stage platform designs to evaluate multiple interventions efficiently. *Clin Trials*. 2022;19(4):432-441. doi:10.1177/17407745221093577

28. Curran D, Sylvester RJ, Hoctin Boes G. Sample size estimation in phase III cancer clinical trials. *European J Surg Oncol*. 1999;25(3):244-250. doi:10.1053/ejso.1998.0635

29. Slud EV. Analysis of factorial survival experiments. *Biometrics*. 1994;50(1):25-38. doi:10.2307/2533194

30. Cro S, Kahan BC, Rehal S, et al. Evaluating how clear the questions being investigated in randomised trials are: systematic review of estimands. *BMJ*. 2022;378:e070146. doi:10.1136/bmj-2022-070146

31. Kahan BC, Morris TP, White IR, Carpenter J, Cro S. Estimands in published protocols of randomised trials: urgent improvement needed. *Trials*. 2021;22(1):686. doi:10.1186/s13063-021-05644-4

32. Mitroiu M, Teerenstra S, Oude Rengerink K, et al. Estimation of treatment effects in short-term depression studies: an evaluation based on the ICH E9(R1) estimands framework. *Pharm Stat*. 2022;21(5):1037-1057. doi:10.1002/pst.2214

33. Tian Z, Esserman D, Tong G, et al. Sample size calculation in hierarchical 2×2 factorial trials with unequal cluster sizes. *Stat Med*. 2022;41:645-664. doi:10.1002/sim.9284

34. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. European Medicines Agency. Published February 17, 2020. Accessed October 18, 2023. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf

35. Pocock SJ, Clayton TC, Stone GW. Challenging issues in clinical trial design: part 4 of a 4-part series on statistics for clinical trials. *J Am Coll Cardiol*. 2015;66(25):2886-2898. doi:10.1016/j.jacc.2015.10.051

36. Leifer ES, Troendle JF, Kolecki A, Follmann DA. Joint testing of overall and simple effects for the two-by-two factorial trial design. *Clin Trials*. 2021;18(5):521-528. doi:10.1177/17407745211014493