

# CONSORT 2025 explanation and elaboration: updated guideline for reporting randomised trials

Sally Hopewell,<sup>1</sup> An-Wen Chan,<sup>2</sup> Gary S Collins,<sup>3</sup> Asbjørn Hróbjartsson,<sup>4,5</sup> David Moher,<sup>6</sup> Kenneth F Schulz,<sup>7</sup> Ruth Tunn,<sup>1</sup> Rakesh Aggarwal,<sup>8</sup> Michael Berkwits,<sup>9</sup> Jesse A Berlin,<sup>10,11</sup> Nita Bhandari,<sup>12</sup> Nancy J Butcher,<sup>13,14</sup> Marion K Campbell,<sup>15</sup> Runcie C W Chidebe,<sup>16,17</sup> Diana Elbourne,<sup>18</sup> Andrew Farmer,<sup>19</sup> Dean A Fergusson,<sup>20</sup> Robert M Golub,<sup>21</sup> Steven N Goodman,<sup>22</sup> Tammy C Hoffmann,<sup>23</sup> John P A Ioannidis,<sup>24</sup> Brennan C Kahan,<sup>25</sup> Rachel L Knowles,<sup>26</sup> Sarah E Lamb,<sup>27</sup> Steff Lewis,<sup>28</sup> Elizabeth Loder,<sup>29,30</sup> Martin Offringa,<sup>13</sup> Philippe Ravaud,<sup>31</sup> Dawn P Richards,<sup>32</sup> Frank W Rockhold,<sup>33</sup> David L Schriger,<sup>34</sup> Nandi L Siegfried,<sup>35</sup> Sophie Staniszewska,<sup>36</sup> Rod S Taylor,<sup>37</sup> Lehana Thabane,<sup>38,39</sup> David Torgerson,<sup>40</sup> Sunita Vohra,<sup>41</sup> Ian R White,<sup>25</sup> Isabelle Boutron<sup>42,43</sup>

For numbered affiliations see end of the article

**Correspondence to:** S Hopewell  
sally.hopewell@csm.ox.ac.uk  
(ORCID 0000-0002-6881-6984)

Additional material is published online only. To view please visit the journal online.

**Cite this as:** BMJ 2025;389:e081124  
<http://dx.doi.org/10.1136/bmj-2024-081124>

**Accepted:** 21 January 2025

Critical appraisal of the quality of randomised trials is possible only if their design, conduct, analysis, and results are completely and accurately reported. Without transparent reporting of the methods and results, readers will not be able to fully evaluate the reliability and validity of trial findings. The CONSORT (Consolidated Standards of Reporting Trials) statement aims to improve the quality of reporting and provides a minimum set of items to be included in a report of a randomised trial. CONSORT was first published in 1996 and was updated in 2001 and 2010. CONSORT comprises a checklist of essential items that should be included in reports of randomised trials and a diagram for documenting the flow of participants through a trial. The CONSORT statement has been updated (CONSORT 2025) to reflect recent methodological advancements and feedback from end users, ensuring that it remains fit for purpose. Here, we

present the updated CONSORT explanation and elaboration document, which has been extensively revised and describes the rationale and scientific background for each CONSORT 2025 checklist item and provides published examples of good reporting. The objective is to enhance the use, understanding, and dissemination of CONSORT 2025 and provide guidance to authors about how to improve the reporting of their trials and ensure trial reports are complete, and transparent.

“Readers should not have to infer what was probably done; they should be told explicitly.” Douglas G Altman<sup>1</sup>

Well designed and properly executed randomised trials provide the most reliable evidence on the benefits of healthcare interventions. Biased results from poorly designed and poorly reported trials are wasteful<sup>2</sup> and can mislead decision making in healthcare at all levels, from treatment decisions for the individual patient to formulation of national public health policies.

Critical appraisal of the quality of randomised trials is possible only if their design, conduct, analysis, and results are completely and accurately reported. However, there is overwhelming evidence that the quality of reporting of randomised trials is not optimal.<sup>3</sup> Without transparent reporting of methods and results, readers cannot evaluate the reliability and validity of trial findings or extract information for systematic reviews. Trials with inadequate methods are also associated with bias, especially exaggerated treatment effects.<sup>4</sup> Having a transparent trial protocol is also important because it prespecifies the design and methods used in the trial, such as the primary outcome, thereby reducing the likelihood of undeclared post hoc changes to the trial.<sup>5</sup>

## SUMMARY POINTS

The CONSORT (Consolidated Standards of Reporting Trials) 2025 statement consists of a 30-item checklist of essential items for reporting the results of randomised trials

This updated explanation and elaboration article describes the rationale and scientific background for each checklist item and provide published examples of good reporting

This explanation and elaboration article provides detailed guidance to enhance the use, understanding, and dissemination of CONSORT 2025, helping to ensure that trial reports are complete and transparent

### **Improving the reporting of randomised trials: the CONSORT statement**

Issues around poor reporting of research are arguably one of the aspects of research waste that is easiest to fix, as highlighted by the late Doug Altman in 1996.<sup>1</sup> Efforts to improve the reporting of randomised trials gathered impetus in the mid-1990s and resulted in the Standardized Reporting of Trials (SORT) statement<sup>6</sup> and Asilomar initiative<sup>7</sup> in 1994. Those initiatives then led to publication of the CONSORT (Consolidated Standards of Reporting Trials) statement in 1996,<sup>8</sup> which was revised in 2001<sup>9</sup> and published alongside an accompanying explanation and elaboration document.<sup>10</sup> CONSORT and the explanation and elaboration document were then updated again in 2010.<sup>11 12</sup> Similar problems related to the lack of transparent reporting of trial protocols led to the subsequent development of the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) statement, published in 2013,<sup>13</sup> and its accompanying explanation and elaboration document.<sup>14</sup> In January 2020, the SPIRIT and CONSORT executive groups met in Oxford, UK. As the SPIRIT and CONSORT reporting guidelines are conceptually linked, with overlapping of content and similar dissemination and implementation strategies, the two groups decided that it was more effective to work together.

The CONSORT statement (or simply CONSORT) comprises a checklist of a minimal set of essential items that should be included in reports of randomised trials and a diagram for documenting the flow of participants through a trial. The objective of CONSORT is to provide guidance and recommendations to authors about how to improve the reporting of their trials and ensure their trial reports are complete, and transparent.<sup>11 12</sup> Readers, clinicians, guideline writers, peer reviewers, and editors can also use CONSORT to help them critically appraise and interpret reports of randomised trials. However, CONSORT is not meant to be used as a quality assessment instrument; it does not specify correct methods to be applied universally. Rather, it reminds the reader of key elements of trial design and conduct related to internal and external validity; when appraising a trial, the reader should critically evaluate the decisions made by the investigators for each of these elements.

Since its publication in 1996, CONSORT has been endorsed by numerous journals worldwide and received global endorsement by prominent editorial organisations, including the World Association of Medical Editors (WAME), International Committee of Medical Journal Editors (ICMJE), and Council of Science Editors (CSE). The endorsement of CONSORT by journals has also been shown to be associated with improved quality of reports of randomised trials.<sup>3</sup>

### **Updating the CONSORT statement**

CONSORT and SPIRIT are living guidelines, and it is vital that the statements are periodically updated to reflect new evidence, methodological advances, and feedback from users; otherwise, their value and usefulness will diminish over time, rendering them

no longer fit for purpose. Updating the SPIRIT 2013 and CONSORT 2010 statements and accompanying explanation and elaboration documents together was also an opportunity to align both checklists and to provide users with consistent guidance in the reporting of trial design, conduct, analysis, and results from trial protocol to final publication. Harmonising the reporting recommendations will improve usability and adherence, leading to more complete and accurate reporting.<sup>15</sup>

The methods used to update the CONSORT statement followed the EQUATOR Network guidance for developers of health research reporting guidelines<sup>16</sup> and have been described in detail elsewhere.<sup>17</sup> In brief, we first conducted a scoping review of the literature related to randomised trials to identify suggestions for changes or additions to CONSORT 2010.<sup>18</sup> We also developed a project specific database (SCEBdb) for empirical and theoretical evidence related to CONSORT and risk of bias in randomised trials.<sup>19</sup> This evidence was combined with evidence from existing CONSORT extensions (Harms,<sup>20</sup> Outcomes,<sup>21</sup> Non-pharmacological Treatment<sup>22</sup>), other related reporting guidelines (Template for Intervention Description and Replication (TIDieR)<sup>23</sup>), and evidence and recommendations from other sources. From these sources, we generated a list of potential modifications or additions to CONSORT, which we presented to end users for feedback in a large international online Delphi survey involving over 300 participants. The Delphi survey results were discussed at a two-day online expert consensus meeting attended by 30 invited international experts. We then drafted the updated CONSORT checklist and revised it based on further feedback from meeting attendees.

The CONSORT 2025 statement is published elsewhere,<sup>24</sup> where full details of the changes to the checklists and rationale for the changes are described. To help facilitate implementation of CONSORT 2025, we have also developed an expanded version of the CONSORT 2025 checklist, with bullet points. The expanded checklist comprises an abridged version of elements presented in this CONSORT 2025 explanation and elaboration document, with examples and references removed (appendix 1).

### **Purpose and main changes to the CONSORT 2025 explanation and elaboration**

We modelled our approach for developing the CONSORT 2025 explanation and elaboration document on the procedures used for the CONSORT 2010 explanation and elaboration document.<sup>12</sup> We present each CONSORT 2025 checklist item (table 1) with at least one recent published example of good reporting, followed by an explanation of the rationale for including the item and main issues to address, and a list selected references to current empirical and theoretical evidence. Members of the core writing group (SH, A-WC, GSC, AH, DM, KFS, IB) developed a draft of individual checklist items, which was then

**Table 1 | CONSORT 2025 checklist of information to include when reporting a randomised trial**

Section/topic	No	CONSORT 2025 checklist item description
<b>Title and abstract</b>		
Title and structured abstract	1a	Identification as a randomised trial
	1b	Structured summary of the trial design, methods, results, and conclusions
<b>Open science</b>		
Trial registration	2	Name of trial registry, identifying number (with URL) and date of registration
Protocol and statistical analysis plan	3	Where the trial protocol and statistical analysis plan can be accessed
Data sharing	4	Where and how the individual de-identified participant data (including data dictionary), statistical code and any other materials can be accessed
Funding and conflicts of interest	5a	Sources of funding and other support (eg, supply of drugs), and role of funders in the design, conduct, analysis, and reporting of the trial
	5b	Financial and other conflicts of interest of the manuscript authors
<b>Introduction</b>		
Background and rationale	6	Scientific background and rationale
Objectives	7	Specific objectives related to benefits and harms
<b>Methods</b>		
Patient and public involvement	8	Details of patient or public involvement in the design, conduct and reporting of the trial
Trial design	9	Description of trial design including type of trial (eg, parallel group, crossover), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)
Changes to trial protocol	10	Important changes to the trial after it commenced including any outcomes or analyses that were not prespecified, with reason
Trial setting	11	Settings (eg, community, hospital) and locations (eg, countries, sites) where the trial was conducted
Eligibility criteria	12a	Eligibility criteria for participants
	12b	If applicable, eligibility criteria for sites and for individuals delivering the interventions (eg, surgeons, physiotherapists)
Intervention and comparator	13	Intervention and comparator with sufficient details to allow replication. If relevant, where additional materials describing the intervention and comparator (eg, intervention manual) can be accessed
Outcomes	14	Prespecified primary and secondary outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome
Harms	15	How harms were defined and assessed (eg, systematically, non-systematically)
Sample size	16a	How sample size was determined, including all assumptions supporting the sample size calculation
	16b	Explanation of any interim analyses and stopping guidelines
<b>Randomisation:</b>		
Sequence generation	17a	Who generated the random allocation sequence and the method used
	17b	Type of randomisation and details of any restriction (eg, stratification, blocking, and block size)
Allocation concealment mechanism	18	Mechanism used to implement the random allocation sequence (eg, central computer/telephone; sequentially numbered, opaque, sealed containers), describing any steps to conceal the sequence until interventions were assigned
Implementation	19	Whether the personnel who enrolled and those who assigned participants to the interventions had access to the random allocation sequence
Blinding	20a	Who was blinded after assignment to interventions (eg, participants, care providers, outcome assessors, data analysts)
	20b	If blinded, how blinding was achieved and description of the similarity of interventions
Statistical methods	21a	Statistical methods used to compare groups for primary and secondary outcomes, including harms
	21b	Definition of who is included in each analysis (eg, all randomised participants), and in which group
	21c	How missing data were handled in the analysis
	21d	Methods for any additional analyses (eg, subgroup and sensitivity analyses), distinguishing prespecified from post hoc
<b>Results</b>		
Participant flow, including flow diagram	22a	For each group, the numbers of participants who were randomly assigned, received intended intervention, and were analysed for the primary outcome
	22b	For each group, losses and exclusions after randomisation, together with reasons
Recruitment	23a	Dates defining the periods of recruitment and follow-up for outcomes of benefits and harms
	23b	If relevant, why the trial ended or was stopped
Intervention and comparator delivery	24a	Intervention and comparator as they were actually administered (eg, where appropriate, who delivered the intervention/comparator, whether participants adhered, whether they were delivered as intended (fidelity))
	24b	Concomitant care received during the trial for each group
Baseline data	25	A table showing baseline demographic and clinical characteristics for each group
Numbers analysed, outcomes, and estimation	26	For each primary and secondary outcome, by group: <ul style="list-style-type: none"> <li>• the number of participants included in the analysis</li> <li>• the number of participants with available data at the outcome time point</li> <li>• result for each group, and the estimated effect size and its precision (such as 95% confidence interval)</li> <li>• for binary outcomes, presentation of both absolute and relative effect size</li> </ul>
Harms	27	All harms or unintended events in each group
Ancillary analyses	28	Any other analyses performed, including subgroup and sensitivity analyses, distinguishing prespecified from post hoc
<b>Discussion</b>		
Interpretation	29	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence
Limitations	30	Trial limitations, addressing sources of potential bias, imprecision, generalisability, and, if relevant, multiplicity of analyses

comprehensively revised based on comments from all authors.

We have made a number of substantive changes to the CONSORT 2025 explanation and elaboration

document. These reflect changes to the CONSORT 2025 checklist: for example, the addition of new items related to data sharing (item 4); how patients and the public were involved in the design, conduct,

and reporting of the trial (item 8); how harms were assessed (item 15); how missing data were handled (item 21c); and details of how the intervention and comparator were actually administered (item 24a). We deleted the checklist item on generalisability of trial findings, which is now incorporated under trial limitations (item 30).

Some changes to the CONSORT 2025 checklist also reflect the integration of aspects of the CONSORT extensions related to harms,<sup>20</sup> outcomes,<sup>21</sup> and non-pharmacological treatments<sup>22</sup>; and TIDieR.<sup>23</sup> We have reorganised the structure of the CONSORT checklist, including a new section on open science, which includes items that are conceptually linked such as trial registration (item 2); where the trial protocol and statistical analysis plan can be accessed (item 3); sharing of de-identified participant level data and statistical code (item 4); and funding and conflicts of interest (item 5). We have aligned the item wording between the CONSORT and SPIRIT 2025 checklists. We have retained some of the same language in the explanation and elaboration text from the previous CONSORT 2010 explanation and elaboration,<sup>12</sup> where it was felt that no changes were required and no new evidence existed to modify our discussion of the issues. The following sections list checklist items (table 1) with examples and explanations.

### **CONSORT 2025: Title and abstract**

#### **Item 1a: Identification as a randomised trial**

##### *Example*

“Efficacy and Safety of Early Administration of 4-Factor Prothrombin Complex Concentrate in Patients With Trauma at Risk of Massive Transfusion: The PROCOAG Randomized Clinical Trial.”<sup>25</sup>

##### *Explanation*

The ability to identify a report of a randomised trial in a bibliographic database depends to a large extent on how it was indexed. Indexers might not classify

a report as a randomised trial if the authors do not explicitly report this information.<sup>26</sup> To help ensure that a study is appropriately indexed and easily identified, authors should use the word “randomised” in the title to indicate that the participants were randomly assigned to their comparison groups.

#### **Item 1b: Structured summary of the trial design, methods, results, and conclusions**

##### *Explanation*

Transparent and sufficiently detailed abstracts are important because readers often base their assessment of a trial on such information. Some readers use an abstract as a screening tool to decide whether to read the full article. However, not all trials are freely available and some health professionals and other users do not have access to the full trial reports.<sup>27</sup>

A journal abstract should contain sufficient information about a trial to serve as an accurate record of its conduct and findings, providing optimal information about the trial within the space constraints and format of a journal. A properly constructed and written abstract helps individuals to assess quickly the relevance of the findings and aids the retrieval of relevant reports from electronic databases.<sup>28</sup> The abstract should accurately reflect what is included in the full journal article and should not include information that does not appear in the body of the paper. In addition, abstracts should not be a distorted representation of the trial results. Studies comparing information reported in a journal abstract with that reported in the text of the full publication have found claims that are inconsistent with, or missing from, the body of the full article.<sup>29-32</sup> Abstracts are also frequently reported with spin defined as a distorted representation of the study results.<sup>33-35</sup> Authors should avoid selectively reporting only statistically significant secondary outcomes or subgroup analyses. Conversely, omitting important harms from the abstract could seriously mislead interpretation of the trial findings and benefit-to-harms balance that is critical for decision making.<sup>36,37</sup>

An extension to CONSORT 2001 provided a list of essential items that authors should include in a journal (or conference) abstract when reporting the main results of a randomised trial.<sup>38</sup> A systematic review of 10 meta-research studies examined the reporting quality of abstracts of randomised trials and found improvements in reporting following publication of this extension.<sup>39</sup> Table 2 provides a list of essential items to include in an abstract; it is based on the CONSORT for Abstracts extension<sup>40</sup> and has been updated to reflect changes made to the main CONSORT checklist. We strongly recommend the use of structured abstracts for reporting randomised trials. They provide readers with information about the trial under a series of headings pertaining to the design, conduct, analysis and interpretation.<sup>41</sup> Some studies have found that structured abstracts offer greater value and information coverage than the more traditional descriptive abstracts<sup>42,43</sup> and allow readers to find

**Table 2 | Items to include when reporting a randomised trial in a journal abstract**

Item	Description
Objectives	Specific objectives
Trial design	Description of the trial design (eg, parallel group, cluster) and framework (eg, superiority, equivalence, non-inferiority, exploratory)
<b>Methods</b>	
Participants	Eligibility criteria for participants and settings where the trial was conducted
Interventions	Intervention(s) and comparator(s) intended for each group
Outcome	Primary outcome(s)
Randomisation	How participants were allocated to interventions (eg, centralised computer generated randomisation)
Blinding	Who was blinded after assignment to interventions (eg, participants, care providers, outcome assessors)
<b>Results</b>	
Numbers randomised	Number of participants randomised to each group
Numbers analysed	For the primary outcome, number of participants analysed in each group
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision
Harms	Important harms or unintended events for each group
Conclusions	General interpretation of the results
Trial registration	Name of trial register and identifying number
Funding	Sources of funding

information more easily.<sup>44</sup> We recognise that journals have their own structure for reporting abstracts. It is not our intention to suggest changes to these formats, but to recommend what information should be reported.

### **CONSORT 2025: Open science**

Open science practices are increasingly endorsed worldwide. There are many definitions of open science and many different goals of openness may be surmised under this term. The United Nations Educational, Scientific and Cultural Organisation defines open science as “an inclusive construct that combines various movements and practices aiming to make multilingual scientific knowledge openly available, accessible and reusable for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society, and to open the processes of scientific knowledge creation, evaluation and communication to societal actors beyond the traditional scientific community.”<sup>45</sup> Open science practices apply to the design, conduct, documentation, and reporting of randomised trial protocols and completed trials.

Some of the checklist items in this section, such as trial registration, were previously included in both SPIRIT and CONSORT. To give them more prominence they are now included under the Open Science heading. However, data sharing is a new item for CONSORT. Furthermore, while no item is dedicated to authorship, we recommend that authors are transparent on whether and how artificial intelligence tools may have been involved in the writing of the manuscript.<sup>46</sup>

#### **Item 2: Name of trial registry, identifying number (with URL) and date of registration**

##### *Example*

“This study was registered in the Iranian Registry of Clinical Trials under the code IRCT20150531022 498N30: <https://en.irct.ir/trial/41031>. Registered on July 26, 2019.”<sup>47</sup>

##### *Explanation*

The consequences of non-publication of entire trials (ie, publication bias),<sup>48-50</sup> and of selective reporting of outcomes and analyses within trials, have been well documented.<sup>51-53</sup> For almost 40 years, there have been growing calls to address these practices. Today, a ubiquitous view recommends trial registration as the best practice to achieve this goal, and inform policymakers and potential participants about ongoing trials. Registering clinical trials, before any assignment of participants, with unique trial identification numbers and other basic information about the trial so that essential details are made publicly available, is a minimum best practice.<sup>54-57</sup> Serious problems of withholding data<sup>58</sup> led to renewed efforts to ensure registration of randomised trials. The World Health Organization (WHO) states that “the registration of all interventional trials is a scientific, ethical and moral responsibility.”<sup>59</sup>

In September 2004, the ICMJE established a policy that it would only consider trials for publication if they had been registered before the enrolment of the first participant.<sup>60</sup> This policy resulted in a substantial increase in the number of trials being registered.<sup>61</sup> However, some trials are still not registered.<sup>49</sup> The ICMJE gives guidance on acceptable registries (<https://www.icmje.org/about-icmje/faqs/clinical-trials-registration/>) and also accepts registration in WHO primary registries (<https://www.who.int/clinical-trials-registry-platform/network/primary-registries>) and ClinicalTrials.gov. Registers charging a fee to view their content should be avoided to ensure equity of access for everyone, including patients and the public.

The Transparency and Openness Promotion (TOP) guidelines, endorsed and used by thousands of journals, recommends trial registration.<sup>62</sup> In a survey of 168 high impact factor medical journals’ “Instructions to authors” in 2014, 78 journals stated that all recent clinical trials must be registered as a requirement of submission to that journal.<sup>63</sup> A more recent survey in 2019 of surgical journals publishing randomised trials found that 53 of 82 journals mandated prospective registration.<sup>64</sup>

Despite recommendations, and mandates in some jurisdictions for clinical trialists to register their trial and evidence that registration deters selective reporting, this is still not happening universally.<sup>65 66</sup> Authors should provide the name of the registry, the trial’s associated registration number, date of registration and, where possible, the URL for the trial’s registration. We recommend that authors also report whether (or when) the trial results are available on the associated trial register.

Despite the considerable increase in clinical trial registration, there is a strong body of evidence showing the lack of access to trial results.<sup>67-70</sup> The latest version of the Declaration of Helsinki states that “Researchers have a duty to make publicly available the results of their research on human participants and are accountable for the timeliness, completeness, and accuracy of their reports . . . Negative and inconclusive as well as positive results must be published or otherwise made publicly available.” In 2015, WHO published a new statement on the public disclosure of trial results, which requests that “the key outcomes are to be made publicly available within 12 months of study completion by posting to the results section of the primary clinical trial registry. Where a registry is used without a results database available, the results should be posted on a free-to-access, publicly available, searchable institutional website of the Regulatory Sponsor, Funder or Principal Investigator.” Some legislations are also in place in the US, UK, and Europe requesting the posting of trial results on clinical trials registry within 12 months after study completion.<sup>71-73</sup>

Authors should indicate whether the trial results are publicly posted to the trial registry, as a preprint (with URL citation) or as published articles (with citations).

**Item 3: Where the trial protocol and statistical analysis plan can be accessed**

*Example*

“The full trial protocol and the Statistical Analysis Plan can be accessed in the Supplementary Material”.<sup>74</sup> This article and supplementary material are open access.

*Explanation*

A protocol for the complete trial (rather than a protocol of a specific procedure within a trial, such as for the intervention) is important because it prespecifies the methods of the randomised trial, for example, the primary outcome (item 14). Having a protocol provides important context to interpret a trial, implement its findings, and facilitate replication and appraisal of risk of bias. It can also help to restrict the likelihood of undeclared post hoc changes to the trial methods and selective outcome reporting (item 10).<sup>75-77</sup> Elements that are important to address in the protocol for a randomised trial are described in the SPIRIT 2025 statement.<sup>78</sup>

A protocol may either include the full statistical analysis plan or may include a section outlining the main principles while referencing and reporting the full statistical analysis plan as a separate, more detailed document. The statistical analysis plan typically includes details about several aspects of the clinical trial, such as the data analysis plan for the primary outcome(s) and all secondary outcomes. Details about what to include in a statistical analysis plan can be found elsewhere.<sup>79</sup>

The protocol should be signed off by the trial steering committee before the allocation of any participants and data collection. Similarly, the full statistical analysis plan should be signed off by the trial steering committee before the dataset is closed for analysis. This allows transparent documentation of any subsequent changes to either document. In many trials, changes to the protocol and statistical analysis plan may happen after the trial onset for legitimate reasons (eg, in response to challenges that were not anticipated or new evidence). In these cases, each iteration of the protocol and statistical analysis plan should record the changes along with their rationale and timing.

There are several options for authors to consider to ensure their trial protocol, and full statistical analysis plan where applicable, are accessible to interested readers. The protocol and full statistical analysis plan (and their various iterations) can be stored in a repository, such as the Open Science Framework, which is free to use and access for all readers. Openness and accessibility are core elements of open science (see Open science section). Trial protocols and statistical analysis plans can also be published in journals such as *Trials* and *BMJ Open*. Open access publication would ensure that any reader, including patients and the public, can access the document. Trial registration (item 2) will also ensure that a minimum set of trial protocol details are available as part of the trial’s registration (<https://www.who.int/clinical-trials-registry-platform>), but often a registration record leaves large ambiguity about key protocol issues and statistical analyses. The more detailed trial protocol can also be posted on most clinical trial registries.

Ideally, the authors should give access to the protocol signed off by the trial steering committee before the allocation of any participants and data collection with any subsequent changes with their rationale and timing.

When submitting a completed trial report, trial authors can include their protocol as a supplemental document or provide a URL to its location. Such documentation can facilitate peer review and help identify reporting biases.

**Item 4: Where and how the individual de-identified participant data (including data dictionary), statistical code, and any other materials can be accessed**

*Examples*

“All data requests should be submitted to the corresponding author (AR) for consideration as agreed in our publication plan. Access to anonymised data may be granted following review with the Trial Management Group and agreement of the chief investigator (AR).”<sup>80</sup>

“Deidentified data collected and presented in this study, including individual participant data and a data dictionary defining each field in the set, will be made available upon reasonable request after publication of this Article, following approval by regulatory authorities. Data can be requested by contacting the corresponding author.”<sup>81</sup>

*Explanation*

Data and code sharing can take the transparency of trial reporting to a different, more desirable level. Sharing individual de-identified participant data would be helpful in many ways: verifying results and increasing trust; using data more extensively for secondary analyses; and using data for individual patient data meta-analysis (IPD MA). Data sharing is also associated with increased citations<sup>82</sup> (ie, broader dissemination). Some trial groups have worked collaboratively to conduct IPD MA.<sup>83</sup> However, for most randomised trials, data sharing does not happen.<sup>84-88</sup> During the covid-19 pandemic, there were many examples of authors’ intentions to share data that then did not transpire (ie, they did not share their data).<sup>87 89</sup> There is increasing concern that some trials are fraudulent or considered to be so-called zombie trials, which becomes evident only on inspection of the raw data.<sup>90 91</sup> However, even if zombie trials are not as prevalent as feared, genuine trials can have such an important role and high value that it is important to maximise their utility by making them more open. Detailed documentation of sharing plans may help in this direction.<sup>92</sup>

All data sharing should abide by the principle of being as open as possible and as closed as necessary throughout a randomised trial’s life cycle (from SPIRIT to CONSORT). It is important to ensure that all the

appropriate permissions are included on the patient consent forms. Trials cannot share data that are not fully anonymised without the appropriate patient consent, and full anonymisation can be difficult. Care must be taken to share participant data appropriately to maintain confidentiality. Suitable mechanisms must be in place to appropriately de-identify participant data, and data should only be shared in a safe and secure manner that fits with the consent obtained from participants.

Data sharing typically involves sharing: the underlying data generated from the trial's conduct; a data dictionary (ie, structure, content, and meaning of each data variable); and other relevant material(s) used as part of the trial's analysis such as the trial protocol, data management plan, statistical analysis plan, and code used to analyse the data. A trial's data can be shared in a variety of ways, such as via an institutional repository (eg, belonging to the university associated with the trial's coordinating centre) and/or a public-facing repository, or by having a bespoke process to provide data. Often, a data use agreement is necessary, which will, at a minimum: prohibit attempts to reidentify or contact trial participants; address any requirements regarding planned outputs of the proposed research (eg, publication and acknowledgment requirements); and prohibit non-approved uses or further distribution of the data.<sup>93</sup>

In a growing number of jurisdictions, funders such as the National Institute for Health (NIH),<sup>94</sup> in the US and the National Institute for Health and Care Research (NIHR) in the UK, alongside other funders such as the Gates Foundation, now require researchers to share their data and make the results publicly available for anyone to read. Similarly, some journals are also requiring authors to include a data sharing statement as part of the article submission process (eg, *Annals of Internal Medicine*, *The BMJ*, *JAMA Network* journals, *PLoS Medicine*).

The process of signalling how data sharing will be achieved is often contained in a data management plan but may also be found in the trial protocol or statistical analysis plan. More complete details regarding developing a data management plan are beyond the scope of this paper. Such details can be found elsewhere.<sup>95</sup> Authors should provide some description of where these details can be found (eg, name of repository and URL to data, code, and materials). Sharing may also entail embargo periods, and if so, the choice of an embargo should be justified and its length should be stated.<sup>96</sup> If data (or some parts thereof) cannot be shared, the reasons for this should be reported and should be sensible and following ethical principles.

For more complex trials (eg, types of talking therapies, physiotherapy), additional materials to share might include a handbook and/or video to detail the intervention.<sup>93</sup> Often these can be shared much more freely than the data, as there are fewer issues with confidentiality.

#### **Item 5a: Sources of funding and other support (eg, supply of drugs), and role of funders in the design, conduct, analysis, and reporting of the trial**

##### *Examples*

"Grant support was received for the intervention from Plan International and for the research from the Wellcome Trust and Joint United Nations Programme on HIV/AIDS (UNAIDS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript."<sup>97</sup>

"Funding: Merck Sharp and Dohme . . . The study funder had a role in the study design, data collection, data analysis, data interpretation, and writing of the report."<sup>98</sup>

The article also states that "Merck employees LY, SB, and PB were involved in the conceptualisation of the study, formal analysis, the investigation process, development of the methodology, project administration, drafting the manuscript, and had critically reviewed and edited the manuscript."<sup>98</sup>

##### *Explanation*

Reporting the funding source(s), and the exact roles of the trial funders, provides important context for readers of a trial report when ascertaining overall methodological rigor (eg, relevance of the type of comparator intervention and eligibility criteria for patients) and risk of bias (eg, selective reporting of favourable results). The trial report should therefore describe details of all funders and the types of funding, as well as the role of the funder in trial design (ie, protocol development), conduct, data analysis, and reporting (ie, interpretation, manuscript writing, and dissemination of results). This should include whether the funder controlled the final decision regarding any of these aspects of the trial, and any mechanisms introduced to minimise funder influence. If the funder had no direct involvement in the trial, that should be stated.

A randomised trial requires considerable funding, typically from pharmaceutical or device companies (industry funding); or from research councils or other scientific or private foundations, or governmental or non-governmental organisations (non-industry funding).<sup>99</sup> One study of trials conducted between 2010 and 2015 estimated the median cost per phase 3 drug company trial at \$21.4m (£17.21m; €20.7m),<sup>100</sup> with substantial variation. The mean cost of clinical trials funded by the NIHR in the UK, reflecting differences in the research and care infrastructure already funded, was lower, but still sizeable—approximately £1.3m—with considerable variation.<sup>101</sup> The various types of funders differ in their overall agenda, their reasons for funding a trial, and their propensity to influence the trial.

Funding of a trial typically involves direct monetary support, but financial support may also be provided indirectly in the form of free trial drugs, equipment, or services (eg, statistical analysis or use of medical writers).<sup>102</sup> Among the most highly cited clinical trials published in 2019 to 2022, two thirds were funded

by industry sponsors, many of whom also provided industry analysts and coauthors.<sup>103</sup>

Industry funding of trials is associated with conclusions that favour the experimental intervention. A systematic review of 75 methodological studies, comparing industry funded studies with non-industry funded studies (mostly randomised trials), reported that industry funded studies had favourable conclusions more often than non-industry funded studies (risk ratio 1.34; 95% confidence interval (CI) 1.19 to 1.51).<sup>104</sup> Industry funded trials may also report more favourable results (ie, larger estimates of intervention effects) than comparable trials that are funded by non-industry sources. One review of eight published meta-epidemiological studies reported that intervention effects (odds ratios) from industry funded trials were, on average, exaggerated by 5% (95% CI -6% to 15%), although the result was imprecise and consistent with chance findings. However, trials with a high risk of industry funder influence (eg, on trial design, conduct, analysis, and reporting) exaggerated effect estimates by 12% (95% CI 3% to 19%).<sup>105</sup> Undue influence on trials from non-industry funders with a strong interest in a specific trial result has been described,<sup>106</sup> but has been studied much less.

A review of 200 trials published in 2015<sup>107</sup> found that 178 (89%) publications included a funding statement. However, in half of the publications, the role of funder was not reported; in the other half, the reporting was often unclear or incomplete; and undisclosed funding from a for-profit organisation was found in 26 of 54 trials reporting only not-for-profit funding. Another study surveyed authors of 200 trials fully funded by industry and found that funders had been involved in the design of 173 trials (87%), in the data analysis of 146 trials (73%), and in the reporting of 173 trials (87%).<sup>102</sup> No clear consensus exists on a monetary threshold for when funding from a source with conflict of interest becomes problematic. It is also unclear whether commercial funding is less important than the degree and type of funder influence on trial design, conduct, analysis, and reporting.

#### **Item 5b: Financial and other conflicts of interest of the manuscript authors**

##### *Example*

“SYR reports grants from Amgen, Astellas, Daiichi Sankyo, Eisai, Merck, Roche, Zymeworks, Indivumed, MSD, Ono/Bristol Myers Squibb, AstraZeneca, BI, Taiho, Lilly, SN Bioscience. SRF has received honoraria as an invited speaker for Lilly, Eisai, Daiichi Sankyo, MSD, and Ono/Bristol Myers Squibb; has participated on advisory boards for Amgen and Indivumed; and has served as an advisor for Astellas, Daiichi Sankyo, Eisai, LG Biochem, Merck Sharpe Dohme, Ono/Bristol Myers Squibb, and AstraZeneca. D-YO reports grants from AstraZeneca, Novartis, Array, Eli Lilly, Servier, BeiGene, Merck Sharpe Dohme, and Handok; and has participated on a data safety monitoring board or advisory board for AstraZeneca, Novartis, Genentech/

Roche, Merck Serono, Bayer, Taiho, ASLAN, Halozyme, Zymeworks, Bristol Myers Squibb/Celgene, BeiGene, Basilea, Turning Point, Yuhan, Arcus Biosciences, IQVIA, and Merck Sharpe Dohme. M-HR reports research grants from AstraZeneca; consulting fees from Bristol Myers Squibb, Ono, Lilly, Merck Sharpe Dohme, Novartis, Daiichi Sankyo, AstraZeneca, Sanofi, and Astellas; and has received honoraria for lectures, presentations, speakers bureaus, or educational events from Bristol Myers Squibb, Ono, Lilly, Merck Sharpe Dohme, Novartis, Daiichi Sankyo, AstraZeneca, Sanofi, and Astellas . . .

“LY, SB and PB report full-time employment by Merck Sharp and Dohme, a subsidiary of Merck (Rahway, NJ, USA), and stock ownership in Merck. LSW reports consulting fees from Amgen; and has received honoraria for lectures, presentations, speakers bureaus, or educational events from Novartis, Bristol Myers Squibb, Merck Sharpe Dohme, Roche, and Amgen.

“PY, YB, JLee, MGF, JLi, MAL, TC, SQ, SL, and HP declare no competing interests.”<sup>98</sup>

##### *Explanation*

Disclosure of authors’ conflicts of interest provides important context for readers of a trial report when ascertaining the overall methodological rigor of a trial (eg, relevance of the type of comparator intervention and eligibility criteria for patients) and risk of bias (eg, selective reporting of favourable results). Conflicts of interest of all trial manuscript authors should be reported, along with any procedures to reduce the risk of conflicts of interest influencing the trial’s design, conduct, analysis, or reporting.

Conflicts of interest can be defined as “a set of circumstances that creates a risk that professional judgement or actions regarding a primary interest will be unduly influenced by a secondary interest.”<sup>108</sup> In the context of authors of a trial report, conflicts of interest imply a risk that investigators’ personal interests and allegiances, or ties with companies or organisations, have undue influence on the design, conduct, analysis, or reporting of a trial. The concept implies a risk of influence and is not indicative of actual wrongdoing.

Conflicts of interest are most often associated with the drug and device industries. Types of financial ties include salary support or grants; ownership of stock or options; honorariums (eg, for advice, authorship, or public speaking); paid consultancy or service on advisory boards and medical education companies; and receipt of patents or patents pending. An analysis of 200 trials from 2015 reported that 57% of trials had at least one author declaring financial conflicts of interests.<sup>109</sup>

Conflicts of interest may also exist with support from or affiliation with government agencies, charities, and professional and civic organisations. Non-financial conflicts of interest include academic commitments; personal or professional relationships; and political, religious, or other affiliations with special interests or advocacy positions. An analysis of 200 trials found

that 4% of trials had at least one author declaring non-financial conflicts of interest.<sup>109</sup> There is ongoing discussion on the association between a problematic non-financial conflict of interest and a reasonable point of view.<sup>110</sup>

A cross sectional study of 190 randomised trials, published in core clinical journals, found that trials with authors' conflicts of interest had more positive results than trials without. The presence of a financial tie was associated with a positive study outcome (odds ratio 3.23; 95% CI 1.7 to 6.1). This association was also present after adjustment for the study funding source (odds ratio 3.57; 95% CI 1.7 to 7.7).<sup>111</sup>

Although financial conflicts of interest are often declared in trials,<sup>109</sup> the declarations are generally imprecise, and undisclosed conflicts are common.<sup>112</sup> A systematic review of studies comparing financial conflicts of interest declared in medical publications or guidelines (not only randomised trials) with declarations in payment databases (eg, the Open Payments Database) found that the median percentage of authors with "non-concordant" disclosures was 81%.<sup>112</sup> A study including only randomised trials found that 35 (30%) of 115 authors from non-industry funded trials had undisclosed conflicts of interest whereas that was the case for 102 (50%) of 203 authors from industry funded trials.<sup>113</sup> For financial conflicts that cannot be tracked to public databases and for non-financial conflicts, the rate of non-disclosure is unknown but it is likely to be even higher.

### **CONSORT 2025: Introduction**

#### **Item 6: Scientific background and rationale**

##### *Example*

"Most problems with shoulder pain are managed in primary care by physiotherapists and GPs [general practitioners]... Evidence from small, short-term trials suggests that physiotherapist-prescribed exercise is promising. However, a Cochrane review highlighted the insufficient evidence about the treatment's long-term clinical effectiveness and cost-effectiveness. Despite widespread provision, uncertainty exists about which types of exercise and levels of physiotherapy supervision are associated with the best outcomes. This evidence is limited by problems in study design and lack of comparator groups. Progressive resistance training to improve muscular strength, whether supervised or home based, has been identified as a core component of exercise for patients with rotator cuff disorders. Subacromial corticosteroid injections are commonly used to reduce local tissue inflammation and pain. Compared with placebo, corticosteroid injections have short-term benefit in the shoulder, although the longer-term benefits and harms are not known. Corticosteroid injections are being used increasingly in clinical practice alongside physiotherapy for the management of people with rotator cuff disorders; hence justification for investigating corticosteroid injection in the GRASP (Getting it Right: Addressing Shoulder Pain) trial alongside physiotherapist-prescribed exercise."<sup>114</sup>

##### *Explanation*

Typically, the introduction of the trial report consists of free-flowing text, in which authors explain the scientific background and rationale for their trial, and its general outline. The rationale may be explanatory (eg, to assess the possible influence of a drug on renal function under tightly regulated conditions) or pragmatic (eg, to guide practice by comparing the benefits and harms of two treatments in a clinical setting). Authors should report any evidence of the benefits and harms of active interventions included in a trial and should suggest a plausible explanation for how the interventions might work, if this is not obvious. Understanding the rationale or theory underpinning an intervention helps readers to understand which aspects or components are likely to be essential to its efficacy.<sup>23</sup> In addition, authors should justify the choice of comparator(s).<sup>115 116</sup> The choice of the comparator (active or placebo) will influence effect estimates. It could raise ethical concerns if patients could be allocated to a placebo or to a suboptimal treatment while an active treatment has been proven effective. Authors should justify the need for the trial they conducted and show that there was equipoise about the best treatment for the condition in the population being studied.

The Declaration of Helsinki states that biomedical research involving people should be based on a thorough knowledge of the scientific literature.<sup>117</sup> It is unethical to expose humans unnecessarily to the risks of research. Some clinical trials have been shown to have been unnecessary because the question they addressed had been, or could have been, answered by a systematic review of the existing literature.<sup>118-120</sup> Thus, the need for a new trial should be justified in the introduction. Ideally, this justification should include a reference to one or more systematic reviews of previous trials. In the absence of a published systematic review, authors should report and summarise the results of previous relevant trials or note their absence.<sup>121</sup> The percentage of published trial reports that cite a systematic review of pre-existing evidence where one is available has increased over time, but over a quarter still fail to do so.<sup>122</sup>

#### **Item 7: Specific objectives related to benefits and harms**

##### *Example*

"To evaluate whether a structured exercise programme improved functional and health related quality of life outcomes compared with usual care for women at high risk of upper limb disability after breast cancer surgery."<sup>123</sup>

##### *Explanation*

Objectives are the questions that the trial was designed to answer. Adequate reporting of the research question is essential to allow readers to appraise and interpret the trial results. The PICO framework, which requires defining the patient population (P); the experimental intervention (I); the comparator intervention or

condition (C); and the outcome or outcomes (O) of interest, has been proposed to help define the research question. PICO is sometimes styled as PICOTS, to include T (the timeframe) and/or S (the setting).<sup>124</sup>

Treatment decisions require an evaluation of the balance between benefit and harm; however, information about harms is frequently omitted or incompletely reported in published reports of trial results.<sup>125-128</sup> Trials whose primary objective is to evaluate benefits of an intervention may not be powered to detect harms, but authors should still report whether they considered harms outcomes when planning the trial.<sup>20</sup>

Authors should clarify whether the aim is to establish superiority of the experimental intervention, or non-inferiority or equivalence, as compared with the comparator intervention.<sup>129</sup> Authors should also report whether the trial is intended to provide preliminary data (a pilot or feasibility trial),<sup>130</sup> explore pharmacokinetic properties, or generate confirmatory results.

For multi-arm trials, authors should clarify which treatment group comparisons are of interest (eg, A v B; A v C). If authors planned to readjust the objective during the trial (eg, in some platform trials or basket trials<sup>131</sup>), this should be reported. Finally, trials can be designed to study the effect of the experimental intervention under different conditions, often described on a spectrum from ideal conditions (explanatory trial) to standard clinical care conditions (pragmatic trial).<sup>132</sup>

The objectives should be phrased using neutral wording (eg, “to compare the effect of treatment A versus treatment B on outcome X for persons with condition Y”) rather than in terms of a particular direction of effect.<sup>133</sup> The trial objectives should align with what was specified in the trial registry and protocol; any changes to the trial objectives after it commenced should be reported with reasons (item 10).

Recently, some trials have been designed using the estimands framework to define the research question and trial objectives. While the terminology surrounding

#### Box 1: Estimands

Concerns have been raised that the precise research questions that randomised trials set out to answer are often unclear.<sup>134</sup> In particular, there is often ambiguity around how events occurring after randomisation (termed intercurrent events) are handled. Specifying the research question using an estimands framework is increasingly used to improve clarity. Despite calls for estimands to be included in the CONSORT 2025 statement,<sup>134 135</sup> their inclusion did not reach consensus. However, we provide a brief overview of estimands and introduce terminology, so they can be applied and reported if used. A more detailed primer on the estimand framework which provides practical guidance on estimands in studies of healthcare interventions can be found elsewhere.<sup>135</sup>

ICH E9(R1) defines an estimand as “a precise description of the treatment effect reflecting the clinical question posed by a given clinical trial objective.”<sup>136</sup> The estimands framework provides a structured description of the objectives in an attempt to bring clarity in specifying the research question, which can be used to guide the study design, data collection, and statistical analysis methods. In brief, an estimand comprises five key attributes: population, treatment groups, endpoint, summary measure, and handling of intercurrent events (table 3). A separate estimand should be defined for each study outcome, and for some outcomes, more than one estimand may be defined.

**Table 3 | Five key attributes of the estimand framework<sup>135</sup>**

Attribute	Definition
Population	Patients for whom researchers want to estimate the treatment effect
Treatment groups	Different intervention strategies being compared in the treatment effect definition
Endpoint	Outcome for each participant that is used in the treatment effect definition
Summary measure	Method used to summarise and compare the endpoint between treatment conditions (eg, risk ratio, odds ratio)
Handling of intercurrent events	Strategies used to handle each intercurrent event* in the treatment effect definition; different strategies could be used for different types of intercurrent events

\*Intercurrent events are post-baseline events (or post-randomisation events in randomised trials) that affect the interpretation or existence of outcome data. These events frequently affect receipt of treatment (eg, treatment switching or treatment discontinuation) or preclude existence of the outcome (eg, death, if it is not defined as part of the outcome).

The ICH E9(R1) outlines five strategies for handling intercurrent events, which are at the core of the estimand framework (table 4).

**Table 4 | Strategies for handling intercurrent events**

Strategy	Description
Treatment policy	The occurrence of the intercurrent event is considered irrelevant in defining the treatment effect of interest: the value for the outcome of interest is used regardless of whether the intercurrent event occurs
Hypothetical	The treatment effect in a scenario where the intercurrent event did not occur is of interest
Composite	The intercurrent event is incorporated into the outcome definition
While on treatment	The outcome before the occurrence of the intercurrent event is of interest
Principal stratum	The outcome in a subpopulation of patients who would not (or would) experience the intercurrent event is of interest.

Although the terminology surrounding estimands may be new to some investigators, it is expected that defining research questions using the estimands framework will become more widespread. A number of existing reporting guidelines have recently included estimands with the reporting recommendations.<sup>137-139</sup> If the estimands framework has been used to design the trial or the data collection or to inform the statistical analysis (by guiding choice of appropriate methods), then this should be made clear in the manuscript and the methods and results should be reported within the framework.

estimands may be new to some investigators, it is expected that the use of this framework will become more widespread. Box 1 provides more information about the estimands framework and how it is being used.

### **CONSORT 2025: Methods**

#### **Item 8: Details of patient or public involvement in the design, conduct, and reporting of the trial**

##### *Examples*

“A study patient advisory group advised on study design before funding, in study set-up, and during recruitment. They chose the term best current treatment and informed the design of clinic procedures (including how best to reduce the burden of intervention), questionnaire design, and participant information. This group informed protocol modifications in response to low recruitment, and guided interpretation of the findings. Two public contributors were members of the independent trial steering committee.”<sup>140</sup>

“The UK Musculoskeletal Trauma Patient and Public Involvement (PPI) Group co-designed this study. In particular, the group advised on the choice of outcome measures and the follow-up arrangements, which were designed to limit the number of face-to-face hospital visits needed. Subsequently, a patient member from this group became a member of the DRAFFT2 Trial Management Group, overseeing all elements of the set-up and delivery of the trial and the dissemination of the lay summary at completion. Another patient member of the group was also a member of the independent Trial Steering Committee.”<sup>141</sup>

“Office workers, workplace champions, and managers within the target organisation were involved in the study design during the grant application process and the study delivery phase. During the grant application phase, the purpose and design of the study as well as the suggested intervention strategies were presented to two large groups of council employees. As a result of these meetings, the study design included using finger prick blood testing rather than taking venous blood samples, participants receiving feedback on health measures, and incentives for attending follow-up. During the study set-up and delivery, a council employee advisory group met several times and provided advice on delivery of the interventions (feedback showed that workplace champions would not be comfortable delivering the initial education session because of the training and planning time required, so this session was delivered online instead), recruitment processes (feedback was provided on participant documents and recruitment messages and strategies within the council), installation of the height adjustable desk, and troubleshooting. Two council employees were also part of the trial steering committee, which met twice a year during the study”.<sup>142</sup>

“In the context of the pandemic and the need to design the study in a short period, no patients were involved in setting the research questions or the outcome measures, nor were they involved

in developing plans for recruitment, design, or implementation of the study. No patients were asked to advise on interpretation or writing up of results.”<sup>143</sup>

##### *Explanation*

Patient and public involvement (PPI) has been shown to be particularly beneficial in clinical trials.<sup>144</sup> It can help researchers to identify and prioritise research topics and questions; identify relevant outcome measures; boost recruitment and retention; improve trial design and tools; and improve the acceptability of trials.<sup>144 145</sup> PPI can also improve the communication and dissemination of the trial results to participants. Public involvement in other types of health research has been shown to help researchers to engage underserved populations and recruit diverse participant groups.<sup>144</sup> Thus, transparent reporting of PPI is essential to allow readers to appraise the relevance and usefulness of findings to end users and to fully evaluate and understand a trial’s methodology and conduct. If patients and the public were not involved, authors should report this with the reasons why.

PPI in health research entails collaborating or partnering with patients and members of the public to design, conduct, report, interpret, or disseminate research: the research is done by or with patients and the public—rather than done for, at, or about them.<sup>146</sup>

<sup>148</sup> Importantly, this is distinct from including patients or members of the public in a trial as participants. PPI contributors can be people with current or past experience of a health condition; their families, carers, and advocates; members of communities who are target users of an intervention or service; or members of the wider public with a broader perspective. The terminology used differs internationally: for example, such activity is most commonly known as “patient and public involvement” in the UK, whereas “patient engagement” is more common in mainland Europe and North America,<sup>147 148</sup> “community and public engagement” is commonly used in part of Africa, and “consumer involvement” is frequently used in Australia.<sup>149</sup>

The GRIPP (Guidance for Reporting Involvement of Patients and Public) checklist was developed in 2011 with the aim of improving the reporting of PPI activities in health research<sup>150</sup>; followed in 2017 by GRIPP2.<sup>151</sup> The GRIPP2 checklist includes GRIPP2-SF, which is a short form of the GRIPP2 checklist where PPI is the secondary focus of the research.<sup>151</sup> Examples of reporting of PPI involvement could include whether and how patients were involved in the trial objectives, whether patients advised on optimising patient recruitment and retention, and whether and how patients were included in selection of the trial outcomes.

Funding bodies are increasingly encouraging or requiring researchers to include PPI in grant applications,<sup>146 152 153</sup> but mandating of PPI reporting by journals has remained uncommon. In 2014, *The BMJ* introduced a requirement for submitted manuscripts to include a PPI statement,<sup>154</sup> and this was extended

to other BMJ journals from 2018.<sup>155 156</sup> A 2023 study of trials addressing chronic conditions found that approximately 80% of trial reports published in these journals included a PPI subsection, and around 40% of these reported that PPI activities had been conducted.<sup>157</sup> Few other journals have followed suit in mandating reporting of PPI activities.<sup>158-160</sup> In the absence of an explicit requirement, reporting of PPI activities remains infrequent, appearing in an estimated 0-5% of published trial reports.<sup>161-163</sup> There is limited evidence to indicate the extent to which this reflects a lack of PPI activity versus PPI activity being conducted but not reported. In a small survey of authors of trials published in high impact factor journals, only one of 29 respondents reported PPI activities having been included but not mentioned in the published article<sup>163</sup>; while in a survey of authors of pragmatic trials, 47% reported including PPI activities during a trial but less than a quarter of these reported the activity in the corresponding publication.<sup>164</sup>

**Item 9: Description of trial design including type of trial (eg, parallel group, crossover), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)**

*Examples*

“This was a multicenter, stratified (6 to 11 years and 12 to 17 years of age), with imbalanced randomization [2:1], double-blind, placebo-controlled, parallel-group study conducted in the United States (41 sites).”<sup>165</sup>

“This multicentre, pragmatic, superiority randomised trial compared three parallel groups for patients referred to secondary care for treatment of primary frozen shoulder, who were recruited from 35 hospital sites in the UK. Individual participants were randomly assigned with unequal allocation (2:2:1) to arthroscopic capsular release, manipulation under anaesthesia, or early structured physiotherapy, to allow for different effect sizes between groups.”<sup>80</sup>

*Explanation*

The word “design” is often used to refer to all aspects of how a trial is set up, but it also has a narrower interpretation. Many aspects of the broader trial design, including details of randomisation and blinding, are addressed elsewhere in the CONSORT checklist. This item refers to the type of trial (eg, parallel group or crossover) and the conceptual framework (eg, superiority, equivalence, or non-inferiority).

CONSORT 2025 focuses mainly on trials with participants individually randomised to one of two parallel groups. While many published trials have such a design, the main alternative designs are multi-arm parallel,<sup>166</sup> crossover,<sup>167</sup> cluster,<sup>168</sup> and factorial designs,<sup>138</sup> with each having their own specific CONSORT extensions. A detailed review of randomised trials published in PubMed in 2012 showed that of the 1122 trials, 953 (85%) were parallel group; the other main designs were crossover (n=98; 13%) and cluster (n=31; 3%). Of these 1122 trials, 892 (80%) had two

groups, 146 (13%) had three groups, and 61 (5%) had four or more groups.<sup>169</sup>

Most trials are designed to identify the superiority of a new intervention, if it exists, but others are designed to assess non-inferiority or equivalence.<sup>129</sup> It is important that researchers clearly describe the design of their trial, including the unit of randomisation (eg, patient, general practice, lesion). These details should also be included in the abstract (item 1b).

If a less common design is used, authors should explain their choice, especially as such designs may imply the need for a larger sample size or more complex analysis and interpretation. Although most trials use equal randomisation (eg, 1:1 for two groups), it is also helpful to provide the allocation ratio explicitly.

**Item 10: Important changes to the trial after it commenced, including any outcomes or analyses that were not prespecified, with reason**

*Examples*

“The original primary endpoint was all-cause mortality, but, during a masked analysis, the data and safety monitoring board noted that overall mortality was lower than had been predicted and that the study could not be completed with the sample size and power originally planned. The steering committee therefore decided to adopt co-primary endpoints of all-cause mortality (the original primary endpoint), together with all-cause mortality or cardiovascular hospital admissions (the first prespecified secondary endpoint).”<sup>170</sup>

“As described in the published protocol paper, a protocol amendment was made to revise the sample size in response to new information on the minimal clinically important difference of the primary outcome measure, the Patient-Oriented Eczema Measure (POEM). Our original sample size used a POEM score for minimal clinically important difference of 3, which was based on research carried out in secondary care among people with moderate or severe eczema. Fresh evidence, however, suggested that a change in POEM score of 2.1 to 2.9 represents a change likely to be beyond measurement error. A protocol amendment was therefore made to change the target sample size for the trials based on seeking to detect a difference in POEM score of 2.5 points between groups, increasing the target sample size from 200 to 303 for each trial.”<sup>171</sup>

*Explanation*

A protocol for a randomised trial serves as the foundation for planning, conduct, reporting, and appraisal, specifying in detail how the trial will be conducted. There may be deviations from the original protocol, as it is impossible to predict every possible change in circumstances during the course of a trial. Some trials will therefore have important changes to the methods after trial commencement. There are many reasons for deviations from the initial study protocol; for example, changes could be due to external information becoming available from other studies, or internal financial difficulties, or lower than anticipated

recruitment rates. In some trials, an independent data monitoring committee will have as part of its remit the possibility of recommending protocol changes based on seeing unblinded data. Authors should report the nature and timing of protocol changes because changes made at different times (eg, before or after breaking the blinding) might be associated with different risks of bias.

Authors should report all major changes to the trial after it commenced indicating the reason for the changes and when the changes occurred.<sup>172</sup> Such changes might affect the trial methods, such as the randomisation ratio, eligibility criteria, interventions, outcomes, method of analysis or duration of follow-up; or might affect the trial conduct, such as dropping a trial site with poor data quality.<sup>173</sup>

Some trials are set up with a formal adaptive design, which allows pre-planned changes to an ongoing trial without compromising the validity of conclusions.<sup>174</sup> It is therefore essential to distinguish pre-planned changes from unplanned changes that may also occur. Such adaptive trial design modifications are usually to the sample size and number of treatment groups, and can lead to decisions being made more quickly and with more efficient use of resources than would be possible with traditional, non-adaptive parallel group trials. Specific guidance has been developed for reporting trials with a formal adaptive design; authors could consult this for more detailed information.<sup>174</sup>

Most trials record multiple outcomes, with the risk that results will be reported for only a selected subset. Prespecification and reporting of completely defined primary and secondary outcomes for both benefits and harms (item 14) should remove such a risk. In some trials, however, circumstances require a change in the way an outcome is assessed, the designation of outcomes as primary or secondary or even, as in the example above, a switch to a different outcome. For example, there may be external evidence from other trials or systematic reviews suggesting the time point for the primary outcome might not be appropriate; or recruitment or the overall event rate in the trial may be lower than expected.<sup>170</sup> Changing an endpoint based on unblinded data are much more problematic, although may be specified in the context of an adaptive trial design.<sup>174</sup>

Whether the modifications are explicitly part of the trial design or in response to changing circumstances, it is essential that they are fully reported and the reason for the change explained to help the reader interpret the results. Such information is not always reported. A comparison of protocols and publications of 102 randomised trials found that 62% of trial reports had at least one primary outcome that was changed, introduced, or omitted compared with the protocol.<sup>51</sup> Primary outcomes also differed between protocols and publications for 40% of a cohort of 48 trials funded by the Canadian Institutes of Health Research.<sup>175</sup> None of these subsequent 150 trial reports mentioned, let alone explained, changes from the protocol. Similar results from other studies have been reported in a

systematic review of empirical studies, comparing trial registers or protocols to published trial reports.<sup>176</sup>

#### **Item 11: Settings (eg, community, hospital) and locations (eg, countries, sites) where the trial was conducted**

##### *Examples*

"The study was conducted in paediatric diabetes services experienced in the use of CSII [continuous subcutaneous insulin infusion], in nine University and six local hospitals within the NHS in England and Wales."<sup>177</sup>

"Setting: One large health board area with a materially deprived, inner city population in the west of Scotland, United Kingdom. Although described as a single centre, NHS Greater Glasgow and Clyde has the largest health board population (1.2m) in the United Kingdom. It is spread over a wide geographical area including severely materially deprived post-industrial areas as well as some more affluent communities. Maternity booking and antenatal care are provided in both hospital and local healthcare settings. Delivery (13 844 infants in 2013[reference]) takes place in three major maternity hospitals."<sup>178</sup>

##### *Explanation*

Along with the eligibility criteria for participants (item 12a) and the description of the interventions (item 13), information on the settings and locations where the trial was conducted is crucial to enable readers to judge the applicability and generalisability of a trial. Were participants recruited from primary, secondary, or tertiary healthcare, or from the community? Healthcare institutions vary greatly in their organisation, experience, and resources and the baseline risk for the condition under investigation. Other aspects of the setting, including the social, economic, and cultural environment, might also affect a study's external validity.

Authors should report the number and type of settings and describe the care providers involved. They should report the locations in which the study was carried out, including the country, city if applicable, and immediate environment (eg, community, general practice, hospital outpatient clinic, or inpatient unit). In particular, it should be reported whether the trial was carried out in one site (single centre trials) or several sites (multicentre trials). The description of the setting should provide enough information so that readers can judge whether the results of the trial could be relevant to their own setting. The environment in which the trial is conducted might differ considerably from the setting in which the trial's results are later used to guide practice and policy.<sup>179 180</sup> Authors should also report any other information about the settings and locations that could have influenced the observed results.

#### **Item 12a: Eligibility criteria for participants**

##### *Example*

"Patients aged 18 years or older were recruited from 20 UK National Health Service (NHS) trusts. Patients

were eligible if they had a diagnosis of shoulder pain attributable to a rotator cuff disorder (eg, cuff tendonitis, impingement syndrome, tendinopathy, or rotator cuff tear) that had started within the past 6 months. We used the diagnostic criteria set out in the British Elbow and Shoulder Society (BESS) guidelines. Patients were excluded if they had a history of significant shoulder trauma (eg, dislocation, fracture, or full-thickness tear requiring surgery), neurological disease affecting the shoulder, other shoulder conditions (eg, inflammatory arthritis, frozen shoulder, or glenohumeral joint instability), received corticosteroid injection or physiotherapy for shoulder pain in the past 6 months, or were being considered for surgery. Detailed criteria are in the protocol.”<sup>181</sup>

#### *Explanation*

A comprehensive description of the eligibility criteria used to select the trial participants is needed to help readers interpret the study. In particular, all inclusion and exclusion criteria should be reported to judge to whom the results of a trial apply—that is, the trial’s generalisability (applicability) and relevance to clinical or public health practice (item 30).<sup>179</sup> A description of the method of recruitment, such as by referral or self-selection (eg, through advertisements) is also important in this context. Because they are applied before randomisation, eligibility criteria do not affect the internal validity of a trial, but they are central to its external validity.

Typical and widely accepted selection criteria relate to the nature and stage of the condition or disease being studied, the exclusion of persons thought to be particularly vulnerable to harm from the study intervention, and to issues required to ensure that the study satisfies legal and ethical norms. The informed consent of study participants, for example, is typically required in intervention studies. Where relevant, it is important to describe whether sex and/or gender were taken into account in the design of the trial, including whether there was adequate representation of men and women (or diverse genders), and justify the reasons for any exclusion.<sup>182</sup>

Despite their importance, eligibility criteria are often not reported adequately. For example, in an analysis of 283 reports of trials published between 1994 and 2006 in high impact general medical journals, reporting of exclusion criteria was often poor and incomplete: 84% of published trials contained at least one poorly justified exclusion criteria, and in 61% more than a quarter of the trial’s exclusion criteria were poorly justified.<sup>183</sup> Any differences in eligibility criteria between the trial protocol and final publication should also be highlighted and reasons for any discrepancies reported. A review of 52 protocols and 75 subsequent full publications submitted to a German medical ethics committee between 2000 and 2006 identified modifications to the eligibility criteria for 85% of trial publications, with 41% of final publications containing newly added criteria.<sup>184</sup> Similar deficiencies have been found in other studies.<sup>185 186</sup>

**Item 12b: If applicable, eligibility criteria for sites and for individuals delivering the interventions (eg, surgeons, physiotherapists)**

#### *Example*

“All participating centres . . . were major neurosurgical centres, treating large numbers of patients after aneurysmal subarachnoid haemorrhage (SAH), each centre treating between 60 and 200 cases annually . . . Centres had to have expertise in both neurosurgical and endovascular management of ruptured aneurysms. Only accredited neurosurgeons with experience of aneurysm surgery were permitted to manage patients in the trial. Endovascular operators had to have done a minimum of 30 aneurysm treatment procedures, before they were permitted to treat patients in the trial”.<sup>187</sup>

#### *Explanation*

For all types of trials, it is important to define the eligibility criteria for clinical sites, and to consider the characteristics of the treatment providers who will provide both the experimental and comparator interventions.<sup>22 188</sup> Evidence suggests that patient outcomes can be associated with hospital and care provider volume.<sup>188</sup> A systematic review of 135 trials found that 71% observed a positive association between hospital volume and outcomes, and 70% observed an association between care provider volume and outcomes.<sup>189</sup> Different levels of expertise of care providers in each intervention group can bias treatment effect estimates.<sup>190</sup> Furthermore, an intervention might be found to be safe and effective in a trial performed in high volume sites by high volume care providers but could have different results in low volume sites. For example, in an analysis of Medicare National Claim files of 167 208 patients who had undergone coronary stent surgery, patients treated by high volume physicians and at high volume sites experienced better outcomes than those treated by low volume physicians at low volume sites.<sup>191</sup> Similar studies show that in most non-pharmacological trials, care provider expertise and site volume will influence treatment effects.<sup>192-194</sup>

Eligibility criteria for care providers and sites are often poorly reported. A systematic review of randomised trials in surgery found that the setting and the site volume of activity were reported in only 7% and 3% of articles, respectively. Eligibility criteria were reported for care providers in 41% of the articles, and the number of care providers performing the intervention was reported in 32%.<sup>195</sup> A careful description of care providers involved in the trial, as well as details of the sites in which participants were treated, helps readers appraise the risk of bias and the applicability of the results.<sup>22 188</sup> Eligibility criteria for sites typically relate to site volume for the procedure under investigation or similar procedures. Eligibility criteria for care providers might include professional qualifications, years in practice, number of interventions performed, skill as assessed by level of complication when performing the intervention,

and specific training before trial initiation. Exclusion criteria should be justified because they will influence the applicability of the trial results.<sup>188</sup>

**Item 13: Intervention and comparator with sufficient details to allow replication. If relevant, where additional materials describing the intervention and comparator (eg, intervention manual) can be accessed**

*Examples*

“Each sulfadoxine–pyrimethamine course consisted of three tablets containing 500 mg of sulfadoxine and 25 mg of pyrimethamine (unbranded generic sulfadoxine–pyrimethamine, Medopharm, Chennai, India; quality controlled by Durbin, Hayes, UK) given as a single oral dose for 1 day (appendix p 2). Each dihydroartemisinin–piperaquine course was dosed according to the bodyweight of each participant and consisted of three to five tablets containing 40 mg of dihydroartemisinin and 320 mg of piperaquine (Alfasigma, Bologna, Italy), given orally once a day for 3 consecutive days. Each dose of azithromycin consisted of two tablets containing 500 mg (Universal Corporation, Nairobi, Kenya) given orally once daily for 2 consecutive days (cumulative dose of 2 g) at the same time as the first and second daily dose of dihydroartemisinin–piperaquine at enrolment. The placebo tablets were also provided by Universal Corporation and had the same appearance as active azithromycin (appendix p 2). The first daily dose was administered in the study clinic under the direct supervision of the study staff, combined with a slice of dry bread or a biscuit. The daily doses on the second and third days were self-administered at home at approximately the same time of the day and in a similar manner as the first dose taken under observation in the clinic.”<sup>196</sup>

“The experimental group received 6 sessions of standard OMT (osteopathic manipulative treatment), and the control group 6 sessions of sham OMT, each session at 2-week intervals. For both experimental and control groups, each session lasted 45 minutes and consisted of 3 periods: (1) interview focusing on pain location, (2) full osteopathic examination, and (3) intervention consisting of standard or sham OMT. Briefly, in both groups, practitioners assessed 7 anatomical regions for dysfunction (lumbar spine, root of mesentery, diaphragm, and atlantooccipital, sacroiliac, temporomandibular, and talocrural joints) and applied sham OMT to all areas or standard OMT to those that were considered dysfunctional. All health care providers were board-certified nonphysician, nonphysiotherapist osteopathic practitioners (*Répertoire National de la Certification Professionnelle*, niveau 1). They all received a 2-day training according to international standards to deliver both standard and sham OMT. Full descriptions of osteopathic practitioner training and interventions are provided in eAppendices 3 and 4 in Supplement 2. In both groups, pharmacological interventions, nonpharmacological interventions, and spinal surgery

were allowed. Cointerventions were self-reported at 3, 6, and 12 months by use of a standardized checklist (eAppendix 5 in Supplement 2).”<sup>197</sup>

*Explanation*

Complete reporting of the intervention and comparator details is essential to enable readers to understand the study results and adequately translate them to clinical practice. Several studies have shown poor reporting of interventions and comparators in randomised trials.<sup>195 198-206</sup> Authors should describe each intervention thoroughly, including control interventions, or use of placebo procedure.<sup>207 208</sup> The description should provide sufficient detail to allow replication, such as to allow a clinician wanting to use the intervention to know exactly how to administer the intervention/comparator that was evaluated in the trial.<sup>198</sup> Key information includes the different components of the intervention/comparator, how and when it should be administered, the intervention/comparator material (ie, any physical or informational materials used in the intervention/comparator, including those provided to participants or used in intervention/comparator delivery or in training of providers and where it can be accessed (eg, online appendix, URL)); the procedure for tailoring the intervention/comparator to individual participants, and how fidelity<sup>209</sup> (ie, the extent to which the intervention/comparator were implemented as planned in the protocol by care providers) or adherence<sup>209</sup> (ie, the extent to which trial participants implement the intervention/comparator as planned in the protocol) were assessed or enhanced (box 2).

Assessing fidelity and adherence can be complex and vary according to the intervention/comparator (eg, one-off, short term repeated, long term repeated). Various deviations to the protocol can occur. For example, participants might initiate the intervention but then discontinue the intervention completely and permanently after a specific period of time, discontinue the intervention temporarily, reduce the dose, or modify the schedule. If relevant, authors should provide the prespecified definition for classifying participants as being treated as planned or not.

In addition, authors should indicate whether criteria were used to guide intervention/comparator modifications and discontinuations and where applicable describe these criteria. This information could be particularly important to evaluate the risk of bias due to deviations from the intended interventions,<sup>210 211</sup> an important domain of the risk-of-bias tool developed by Cochrane.<sup>210</sup> Assessing this domain requires a clear understanding of, and ability to distinguish between, deviations that occur as planned in the protocol and deviations that arise due to the experimental context.

The research question (ie, explanatory v pragmatic) will affect the standardisation of the intervention/comparator as well as how adherence or fidelity is assessed or enhanced. In explanatory trials, the aim is to estimate treatment effect under ideal circumstances.

**Box 2: Examples of essential information to be reported for various types of interventions in randomised trials\***<sup>22,23</sup>
**Drug<sup>23</sup>**

- Generic name
- Manufacturer
- Dose
- Route of administration (eg, oral, intravenous)
- Timing
- Titration regimen if applicable
- Duration of administration
- Procedure for tailoring the intervention to individual participants
- Conditions under which interventions are withheld
- Whether and how adherence of patients to the intervention was assessed or enhanced
- Any physical or informational materials used in the intervention and where the materials can be accessed.

**Rehabilitation, behavioural treatment, education, and psychotherapy, etc<sup>22,23</sup>**

- Qualitative information
  - Theory/rationale for essential intervention elements
  - Content of each session
  - Mode of delivery (individual/group, face to face/remote)
  - Whether the treatment is supervised
  - The content of the information exchanged with participants
  - The materials used to give information
  - Procedure for tailoring the intervention to individual participants
  - Whether and how the interventions were standardised
  - Background and expertise of individuals delivering the interventions
  - Whether the same care providers delivered interventions across trial groups
  - Whether and how adherence of individuals delivering the interventions to the protocol was assessed or enhanced
  - Whether and how adherence of patients to the intervention protocol was assessed and/or enhanced
  - Any physical or informational materials used in the intervention and where the materials can be accessed.
- Quantitative information
  - Intensity of the intervention where appropriate
  - Number of sessions
  - Session schedule
  - Session duration
  - Duration of each main component of each session
  - Overall duration of the intervention.

**Surgery, technical procedure, or implantable device<sup>22</sup>**

- Preoperative care relevant details
- Intraoperative care relevant details
- Configuration of any device
- Postoperative care relevant details
- Procedure for tailoring the intervention to individual participants
- Whether and how the interventions were standardised
- Background and expertise of individuals delivering the interventions
- Whether the same care providers delivered interventions across trial groups
- Whether and how adherence of individuals delivering the interventions to the protocol was assessed or enhanced
- Any physical or informational materials used in the intervention and where the materials can be accessed.

\*This list is not intended to be exhaustive; it is a starting point for authors to consider when reporting the intervention.

The intervention/comparator are usually highly standardised with close monitoring of fidelity and adherence to the intervention/comparator and strategies to increase them. In contrast, pragmatic trials aim to determine treatment effect in clinical conditions. The intervention and comparator are usually highly flexible, and measurement of fidelity and adherence are unobtrusive with no strategies to maintain or improve them. Nevertheless, assessing

fidelity and adherence to the intervention/comparator, or at least recording the most important components of the intervention/comparator, is necessary to understand what was actually administered to participants. This is particularly important for complex interventions where diversity in the implementation of the intervention is expected. For example, in a pragmatic trial assessing a surgical procedure where the procedure is left to surgeons' choice, investigators

should plan to systematically record key elements related to pre-operative care, anaesthesia, the surgical approach, and post-operative care. This information is essential to provide a relevant description of the intervention that was actually provided when the trial is completed.

If the control group or intervention group received a combination of interventions, the authors should provide a thorough description of each intervention, an explanation of the order in which the combination of interventions were planned to be introduced or withdrawn, and the triggers for their introduction if applicable. Some complex interventions will require the development of specific documentation (eg, training materials, intervention manuals). Authors should make these available and indicate where they can be accessed.

If the control group is to receive usual care, it is important to describe what that constitutes so that readers can assess whether the comparator differs substantially from usual care in their own setting.<sup>212</sup> Various approaches could be used: standardising usual care to be in line with specific guidelines; or asking practitioners to treat control patients according to their own preference, which could result in heterogeneity of the care provided particularly between centres and over time.<sup>213</sup> Usual care can vary substantially across sites and patients, as well as over the duration of the trial. Further, it is important to clarify if the experimental group also received usual care in addition to the experimental intervention, and what actually differed between the groups. Usual care is frequently incompletely reported. In a review of 214 paediatric trials, the descriptions of standard of care were more often incomplete than the description of the intervention arms within the same study as measured by the TIDieR checklist (mean 5.81 (standard deviation (SD) 2.13) v 8.45 (SD 1.39)).<sup>205</sup>

If the control group is to receive a placebo, specific considerations need to be accounted for. Some evidence showed that placebos are insufficiently described.<sup>214</sup> Placebo could have several different aspects from pills to saline injections or more complex interventions such as sham interventions (eg, sham surgery) or attention control interventions. Authors should report the same level of details as required for the intervention—that is, content of the placebo or different component of the placebo, how and when it should be administered, material, procedure for tailoring the placebo to individual participants, and how fidelity and adherence were assessed or enhanced.<sup>207</sup> Complete reporting of the placebo is needed to understand what intervention effect is measured in the trial.<sup>215</sup> A network meta-analysis of osteoarthritis trials showed that different placebo interventions (oral, intra-articular, topical, oral and topical) had different effects and can impact the relative effect estimate of active treatments.<sup>216</sup>

Further, the trial groups could receive different concomitant care in addition to the assigned trial interventions. Concomitant care could impact trial

outcomes and bias effect estimates. To facilitate interpretation of study results and risk-of-bias assessments, authors should report relevant concomitant care that was allowed or prohibited where relevant.

Specific guidance has been developed to improve the reporting of interventions, particularly TIDieR,<sup>23</sup> TIDieR-Placebo for placebo and sham controls,<sup>207</sup> and the CONSORT extensions for non-pharmacological treatments.<sup>22</sup> Authors could consult these for more detailed information.

**Item 14: Prespecified primary and secondary outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome**

*Example*

See table 5.<sup>21</sup>

*Explanation*

All randomised trials assess outcomes, for which the groups are compared. Most trials have several outcomes, some of which are of more importance than others. The primary outcome is the prespecified outcome considered to be of greatest importance to relevant stakeholders (such as patients, policy makers, clinicians, and funders) and should be the one used in the sample size calculation (item 16). The primary outcome should be explicitly indicated as such in the report of a randomised trial. Other outcomes of interest are secondary outcomes.

It is important to explain the rationale and clinical relevance for chosen efficacy and harm outcomes, including whether they are part of a core outcome set.<sup>217 218</sup> A core outcome set is an agreed standardised set of outcomes that should be measured and reported, as a minimum, in all clinical trials in specific areas of health or health care. The COMET (Core Outcome Measures in Effectiveness Trials) initiative and COMET database facilitate access to core outcome sets (<https://www.comet-initiative.org/>).

Most trials have a single primary outcome. Having several primary outcomes can incur potential problems of interpretation associated with multiplicity of analyses (items 28 and 30). There are typically multiple secondary outcomes (ie, the outcomes prespecified in the trial protocol to assess any additional effects of the intervention). Secondary outcomes can include harms that may include unintended effects of the intervention (item 27).

The primary and secondary outcomes reported should be consistent with the outcomes prespecified in the trial protocol and the registry. Evidence shows important discrepancies between the outcome reported in the registry or protocol and outcomes reported in final publications, frequently in favour of statistically significant results. Any change in outcome(s) specified in the protocol should be reported, with reasons (item 10).<sup>51 175 176</sup>

**Table 5 | Five core elements of a defined outcome with examples**

Element term	Definition used	Example 1	Example 2	Example 3
Domain	Title or concept to describe one or more outcomes	Blood pressure	Depression	Death
Measurement variable or specific measurement	Corresponds to the data collected directly from the trial participants; description includes the instrument used to assess the outcome domain	—	—	—
	• Descriptive name	Systolic blood pressure measured with Omran upper arm blood pressure monitor	MADRS	All cause mortality, per hospital database
	• If applicable, the total score or subscale that will be analysed	Not applicable	MADRS total score	Not applicable
Specific metric	Participant level unit of measurement (eg, change from baseline, final value or a value at a time point, time to event) for the analysis	Value at a time point	Change from baseline	Time to event
Method of aggregation	Procedure for estimating the treatment effect	—	—	—
	• If the outcome will be treated as continuous, categorical, or, time to event variable	Continuous variable	Binary variable	Time to event
	• For continuous variables, a measure of central tendency (eg, mean value); for categorical and time-to-event data variables, proportion with an event, and (if relevant) the specific cut-off values or categories compared	Mean value	Proportion of participants with ≥50% decrease	Incidence density and between group incidence density rate
Time point	Timing of follow-up measurements	—	—	—
	• When outcome measurements will be obtained	2, 4, and 12 weeks after randomisation	2, 4, 6, and 8 weeks after randomisation	Daily
	• Which of the outcome measurements will be analysed	12 weeks after randomisation	8 weeks after randomisation	End of follow-up

MADRS=Montgomery-Asberg Depression Rating Scale. Table adapted from Butcher et al.<sup>21</sup>

All outcomes, whether primary or secondary, should be described and completely defined. This information is typically also detailed in the trial's protocol and the trial registry. The principle here is that the information provided should be sufficient to allow others to use the same outcomes.<sup>198</sup> For each outcome, it is important to detail: (1) the specific measurement variable, which corresponds to the data collected directly from trial participants (eg, Beck Depression Scale; all cause mortality) with definition where relevant (eg, major bleeding was defined as fatal bleeding or symptomatic bleeding in a critical area or organ; all cause mortality per hospital database); (2) the specific participant level analysis metric, which corresponds to the format of the outcome data that was used from each trial participant for analysis (eg, change from baseline; final value or value at a time point; time to event); (3) the method of aggregation, which refers to the summary measure format for each trial group (eg, mean; proportion of participants with score >2); and (4) the measurement time point of interest for analysis.<sup>219</sup> For composite outcomes, all individual components of the composite outcome should be described as secondary outcomes.<sup>21</sup> Only half of randomised trials published in PubMed indexed journals in 2000 and 2006 specified the primary outcome.<sup>220 221</sup> In recent samples of trials published in specific fields, reporting has improved but still two thirds did not provide a complete definition.<sup>222 223</sup>

The use of previously developed and validated scales can help to enhance quality of measurement.<sup>224 225</sup> For example, assessment of health related quality of life using a validated instrument is critical to the integrity and applicability of the trial.<sup>226</sup> Authors should report

measurement properties of outcome measurement instruments to assist in interpretation and comparison with similar studies.<sup>227</sup>

In most trials, information on outcomes is set to be collected as part of the trial conduct. However, some trials may use existing data collecting structures (eg, national, healthcare or administrative registries). This should be clarified in the methods. There is empirical evidence that treatment effect estimates may be different in trials where outcomes are obtained from routinely collected data.<sup>228</sup>

#### Item 15: How harms were defined and assessed (eg, systematically, non-systematically)

##### Examples

"Adverse events (AE) were assessed clinically and analytically at each monthly follow-up visit. The severity of AE were classified according to the National Cancer Institute Common Toxicity Criteria version 4.0. Following the onset of the first cases, the criteria for considering the presence of tenosynovitis were established as spontaneous pain that increased with movement in any tendon insertion with tenderness at that level and observation of localized inflammatory signs of at least 72 hours in duration... Patients were considered to have hepatotoxicity when they presented alanine transaminase, aspartate aminotransferase, or bilirubin elevations more than 2 times the upper limit of the normal range. Toxicity was considered severe, and therefore the drug was discontinued when symptomatic elevations were more than 3 times or asymptomatic elevations were more than 5 times the normal levels. All AE were recorded and additional information was required in case of serious adverse events."<sup>229</sup>

"Immediate adverse events were assessed by monitoring participants 30 min after injection in the trial centre. All participants were required to report all local and systemic adverse reactions and adverse events after the injection using the trial's mobile application. Solicited adverse reactions were defined as any events that occurred from day zero to day seven after each injection. Unsolicited adverse reactions were defined as any adverse reactions which occurred from day eight to day 28 after each injection. The severity of adverse reactions was defined using the Food and Drug Administration guidance for industry and toxicity grading scale for healthy adult and adolescent volunteers enrolled in preventive vaccine clinical trial."<sup>230</sup>

#### *Explanation*

Evaluation and reporting of harms in randomised trials can be useful to inform decision makers on the benefit-risk balance of an intervention.<sup>20-36</sup> Randomised trials usually lack power and sufficient follow-up to adequately estimate harms<sup>231</sup>; nevertheless, they can provide data about harms that can be synthesised in meta-analyses if adequately reported.<sup>20,36</sup> For example, the Women's Health Initiative trials on hormone therapy provided important data on the cardiovascular risk of hormone replacement therapy.<sup>232</sup>

Harm relates to the unwanted effects of an intervention. According to the specific context, a given event could be considered for assessing harm (eg, myocardial infarction in a trial assessing non-steroidal anti-inflammatory drugs in patients with

osteoarthritis) or benefit (eg, myocardial infarction in a trial assessing aspirin in patients with cardiovascular risk factors) of an intervention.<sup>232</sup> The use of the term "harms" is preferred over "safety" to better reflect the negative effect of interventions.<sup>20</sup>

Despite the importance of having access to data on harms, reporting of this information is poor.<sup>127,128,233,234</sup>

A review of 184 drug trials published between 2015 and 2016 in four medical journals with high impact factors showed that 28% did not provide any details on how harm data were collected and 89% did not report who decided whether the harm was attributable to the study drug.<sup>128</sup> An overview of 13 reviews assessing the reporting of harms in randomised trials using the 2004 CONSORT extension for harms showed that only 40% of the included trials addressed harms outcomes with definitions for each and 44% clarified how harms-related information was collected.<sup>127</sup>

How harms are defined and assessed will affect the results and effect estimates. Harm can be prespecified or not. They can be systematically assessed or rely on spontaneous declaration (ie, non-systematically assessed). To increase the study's power, harm can be aggregated in a composite outcome (eg, cardiovascular diseases). Some trials implement a procedure to determine whether harms could be attributed to the intervention (ie, causality).

For each systematically assessed harm, authors should report the definition, the measurement variable (eg, name of a validated questionnaire), and where appropriate, the analysis metric for each participant (eg, time to event), the summary measure for each trial group (eg, proportion), and the time point of interest for analysis. They should describe the procedures for harm assessment including who did the assessment, whether they were blinded to the treatment allocated, the assessment time points, and the overall time period for recording harms. For non-systematically assessed harms, authors should report the mode of data collection with the time point and overall time period for recording harms. Nevertheless, non-systematically assessed harms can be difficult to analyse and interpretation of the results should be viewed with caution. To overcome this issue, trialists can code and group events into specific categories. Nevertheless, the lack of standardisation in data collection could result in selective and incomplete reporting.<sup>231</sup> Access to individual participant data may be needed to adequately synthesise this information.<sup>231,235,236</sup>

Where appropriate, the process for coding each harm and grading its severity should be described including who did the coding and severity grading, and whether they were blinded to the allocated trial group.

If harms outcomes are aggregated (eg, cardiovascular events, serious events, severe events, withdrawals due to harms, harms imputed to treatment), authors should describe the process for classifying harms, including the grouping system used (eg, grading system to define severity), who did the grouping, and whether they were blinded to the treatment allocated.<sup>237,238</sup>

### **Box 3: Reporting harms in randomised controlled trials**

#### **Systematically assessed harms**

- Definition and instrument used (eg, name of a validated questionnaire)
- Analysis metric for each participant (eg, time to event); summary measure for each trial group (eg, proportion); time point of interest for analysis, where appropriate.

#### **Procedures for harm assessment, including:**

- Who did the assessment
- Whether the assessors were blinded to the allocated trial group
- Assessment time points and the overall time period for recording harms

#### **Non-systematically assessed harms**

- How data were collected
- Assessment time points and overall time period for recording harms

#### **Coding harms and grading severity**

Where appropriate, process for coding each harm and grading its severity, including:

- Who did the coding and severity grading, and whether they were blinded to the allocated trial group
- Which coding and severity grading systems were used, if any
- Assessment time points and overall time period for recording harms

#### **Grouping of harms**

For grouping of harms by body system, seriousness, severity, withdrawals (due to harms), and causality:

- Definitions of grouping categories
- Who did the grouping, and whether they were blinded to the allocated trial group.

Box 3 summarises the essential information to be reported related to harms. More detailed information can be found in the CONSORT extension for harms, which was updated in 2022.<sup>20</sup>

**Item 16a: How sample size was determined, including all assumptions supporting the sample size calculation**

*Examples*

“We expected an improvement in PFS [progression free survival], in favor of avelumab, with a hazard ratio (HR) of 0.58. Considering a fixed design with a 2-sided  $\alpha$  risk of 5% and a power of 80%, 106 events (progression or death) are needed to demonstrate this difference based on the Schoenfeld method. With an estimated recruitment rate of 3 patients per month, a follow-up period for each patient of 24 months, and a percentage of patients lost to follow-up or not evaluable of 15%, 132 patients had to be randomized, and we planned to enroll a total of 66 patients per group.”<sup>239</sup>

“The target sample size was 300 (150 per arm) over a 3-and-a-half-year recruitment period. This was based on an assumed proportion of individuals with clinically meaningful improvement in VA [visual acuity] (>10 letters) of 55% in the standard care arm and a 19% increase in the adjunct group to 75%, with approximately 7% loss to follow-up, at least 90% power and two-sided 5% type 1 error.”<sup>240</sup>

“In order to detect a minimum clinically important difference (MCID) in mean volume of daily PA [physical activity] of 2.1 m g [milligravity] at 12 months, and assuming a standard deviation (SD) of 5.3 m g, power of 80%, and a statistical significance level of 5%, a total of 202 participants were required. Allowing for 20% loss to follow-up and 20% non-compliance of accelerometer/intervention attendance meant that at least 338 participants were required (169 per group). The value of 2.1 m g was chosen as it represents an increase in PA that is equivalent to walking at the threshold between light intensity and moderate intensity (for example, 4 km per hour) for 30 min per day or 10–15 min of brisk walking per day.”<sup>241</sup>

“Sample size was based on the primary outcome measure, HOS ADL [hip outcome score activities of daily living subscale] at eight months post-randomisation, and was calculated using a minimum clinically important difference between groups of 9 points. We estimated the standard deviation to be 14 points; however, summaries presented at a planned interim data monitoring meeting found that the standard deviation was 18 points. A revised calculation (significance level 5%, power 90%, loss to follow-up 20%) gave a sample size of 214 (107 participants in each group). The data monitoring committee approved the sample size increase from 120 to 214 participants.”<sup>242</sup>

*Explanation*

Sample size calculations are a key design component for a trial and need careful planning. Sample size calculations need to balance ethical and logistical

considerations alongside medical and statistical considerations so that the scientific question can be reliably and precisely answered in a timely manner without unnecessarily exposing individuals to ineffective or harmful interventions. They are generally based on one primary outcome. A trial should therefore be sufficiently large to have a high probability (power) of identifying a clinically important difference of a prespecified size that meets a criterion of statistical significance, if such a difference exists. The magnitude of the effect has an inverse relationship with the sample size required for its detection; that is, larger sample sizes are needed to detect smaller differences. Moreover, the inverse relationship is not linear: very small differences require enormous sample sizes to have good power to detect.

All details on how the sample size was determined should be reported to allow replication (in principle). Elements of the sample size calculation that need to be specified are the primary outcome (and time point) on which the calculation was based (item 14); the anticipated values for the outcome in each trial group (which implies the clinically important target difference between the intervention groups) at a specific time point with rationale or provenance of all quantities, including any relevant citations; or continuous outcomes, the standard deviation of the measurements<sup>243</sup>; the statistical test; the  $\alpha$  (type I error) value and whether it is two sided; the statistical power (or the  $\beta$  (type II error) value); and the resulting target sample size per trial group (box 4). Details should be given of any inflation of the sample size made for attrition or non-adherence during the study. Reference to any formulas or software packages used for the sample size calculation should all be reported. The reporting will have additional considerations for crossover trials,<sup>167</sup> factorial trials,<sup>138</sup> cluster trials,<sup>168</sup> multi-arm trials,<sup>166</sup> within-person trials,<sup>245</sup> and non-inferiority and equivalence trials.<sup>129</sup>

Transparency in the sample size reveals the power of the trial to readers and gives them a measure by which to assess whether the trial attained its planned size. Any differences in the planned sample size described in the trial registration (item 2), study protocol (item 3), or statistical analysis plan should be explained.

Interim analyses are used in some trials to help decide whether to stop early or to continue recruiting sometimes beyond the planned trial end (item 16b). If the actual sample size differed from the originally intended sample size for some other reason (eg, because of poor recruitment or revision of the target sample size), an explanation should be given alongside details of the revised sample size. Many reviews have found that few authors report how they determined the sample size.<sup>222 246-252</sup>

There is no value in conducting and reporting a post hoc calculation of statistical power using the results of a trial, for example, as a pretext to explain non-significant findings; this may even mislead and confuse readers.<sup>253 254</sup>

**Box 4: DELTA<sup>2</sup> recommended reporting items for the sample size calculation of a randomised controlled trial with a superiority question\***<sup>244</sup>

**Core items**

1. Primary outcome (and any other outcome on which the calculation is based). If a primary outcome is not used as the basis for the sample size calculation, state why
2. Statistical significance level and power
3. Express the target difference according to outcome type
  - a. Binary—state the target difference as an absolute or relative effect (or both), along with the intervention and control group proportions. If both an absolute and a relative difference are provided, clarify if either takes primacy in terms of the sample size calculation
  - b. Continuous—state the target mean difference on the natural scale, common standard deviation, and standardised effect size (mean difference divided by the standard deviation)
  - c. Time to event—state the target difference as an absolute or relative difference (or both); provide the control group event proportion, planned length of follow-up, intervention and control group survival distributions, and accrual time (if assumptions regarding these values are made). If both an absolute and relative difference are provided for a particular time point, clarify if either takes primacy in terms of the sample size calculation
4. Allocation ratio. If an unequal ratio is used, the reason for this should be stated
5. Sample size based on the assumptions as per above
  - a. Reference the formula/sample size calculation approach, if standard binary, continuous, or survival outcome formulas are not used. For a time-to-event outcome, the number of events required should be stated
  - b. If any adjustments (eg, allowance for loss to follow-up, multiple testing) that alter the required sample size are incorporated, they should also be specified, referenced, and justified along with the final sample size
  - c. For alternative designs, additional input should be stated and justified
  - d. Provide details of any assessment of the sensitivity of the sample size to the inputs used

**Additional items for grant application and trial protocol**

1. Underlying basis used for specifying the target difference (an important or realistic difference)
2. Explain the choice of target difference—specify and reference any formal method used or relevant previous research

**Additional item for trial results paper**

1. Reference the trial protocol.

\*Taken from Cook et al.<sup>244</sup>

**Item 16b: Explanation of any interim analyses and stopping guidelines**

*Examples*

“Interim analyses of effectiveness and safety endpoints were performed on behalf of the data monitoring committee on an approximately annual basis during the period of recruitment. These analyses were done with the use of the Haybittle–Peto principle and hence no adjustment was made in the final p values to determine significance.”<sup>255</sup>

“One interim analysis of the primary endpoint and safety data was planned for when approximately 50% of the participants had completed D28 [day 28]. Statistical significance and futility boundaries were estimated for the interim and final analysis based on 50,000 simulations from the PASS® software (NCSS, Kaysville, Utah) by simulating a group sequential test for two means assuming normality testing. At the interim analysis, the two-sided significance boundary for clinical efficacy was 0.00312 and for futility of detecting  $\mu_{AUT00063} > \mu_{placebo}$ , the one-sided O’Brien–Fleming boundary was 0.39,141. Hence, at the final analysis, the two-sided significance boundary for clinical efficacy would be 0.04761. The Independent Data Monitoring Committee (IDMC) was advised to consider making recommendations

for early termination only where there was a clear demonstration of futility.”<sup>256</sup>

“Three planned analyses (two interim analyses and one final analysis) were performed when the observed number of events were 25, 47, and 84, respectively. Data were released by DSMC [data and safety monitoring committee] after final analysis. Efficacy stopping boundaries were based on the O’Brien–Fleming spending function. Futility boundaries were based on testing the alternative hypothesis at the 0.039 level.”<sup>257</sup>

“Two interim analyses to be performed using the Haybittle–Peto approach were scheduled, after enrolment of 1000 and 2000 patients, respectively. The significance level associated with both interim analyses was 0.001 and the significance level associated with the final analysis was 0.049. With this method, the overall risk of type 1 error was 5%.”<sup>258</sup>

*Explanation*

Numerous randomised trials enrol participants over extended periods of time. If an intervention demonstrates exceptional efficacy, the study might require early termination on ethical grounds. To mitigate this concern, assessing results as data accumulates is advisable, ideally through an independent data

monitoring committee (DMC), sometimes referred to as a data and safety monitoring board (DSMB).<sup>259</sup> However, conducting multiple statistical evaluations on accruing data without proper adjustment may result in misleading conclusions. For instance, examining data from a trial at five interim analyses using a P value of 0.05 would elevate the overall false-positive rate closer to 19% rather than the expected 5%. Further to stopping early for efficacy, interim analyses can be used to evaluate (1) futility, to assess whether a trial is likely to meet its objectives; or (2) safety, to assess whether there is evidence for increased risk of harms (in the intervention group relative to the comparator group).<sup>259</sup> Interim analyses can also be used to reassess the sample size, using updated information from interim trial data (eg, through an internal pilot), to ensure adequate power of the trial.

Various group sequential statistical approaches exist to adjust for multiple looks (ie, analyses) at the data, and these should be predetermined in the trial protocol (see item 27a of the SPIRIT 2025 statement<sup>78</sup>). Using these methods, data are compared at each interim analysis, where a P value below the specified critical value by the chosen group sequential method signifies statistical significance. Some researchers view group sequential methods as a tool for decision making, while others regard them as a definitive stopping point, intending to halt the trial if the observed P value falls below the critical threshold.

Authors should disclose whether they or the DMC/DSMB performed multiple looks at the data (interim analyses). If such multiple looks occurred, it is important to specify the frequency; the triggers prompting them; the statistical methods applied

(including any formal stopping rules); and whether these procedures were planned and documented in the trial protocol before the trial commenced, before the DMC examined any interim data, or at a later stage. Authors should also report the time point at which any interim analyses were conducted (and by whom); and state who decided to continue, stop or modify the trial, and whether they were blinded to the treatment allocation. Unfortunately, the reporting of interim analyses and stopping rules is frequently inadequate in published trial reports,<sup>260</sup> even in cases where trials indeed halted earlier than originally planned.

#### **Item 17a: Who generated the random allocation sequence and the method used**

##### *Examples*

“Randomization was done using computer-generated random numbers (Stat Trek software) by trained staff at the Soltan Mirahmad Clinic (Kashan, Iran).”<sup>261</sup>

“The randomization was conducted by two independent researchers who were not involved in the study using a computer random sequence generator.”<sup>262</sup>

##### *Explanation*

Randomisation eliminates selection bias at trial entry and is the crucial component of high quality randomised trials (box 5).<sup>271</sup> Successful randomisation hinges on two steps: generation of an unpredictable allocation sequence and concealment of this sequence from the investigators enrolling participants (item 18).<sup>268 269</sup>

Who generated the random allocation sequence is important mainly for two reasons. Firstly, someone, or some group, should take responsibility for this critical trial function. Secondly, providing information on the generator might help readers to evaluate whether anyone had access to the allocation sequence during implementation. Investigators should strive for complete separation, independence, between the trial staff involved with generation of the allocation sequence and those staff who implement assignments (see explanation for item 19).

Participants should be assigned to comparison groups in the trial on the basis of a chance (random) process characterised by unpredictability (box 5). Successful randomisation in practice depends on two inter-related aspects: adequate generation of an unpredictable allocation sequence and concealment of that sequence until assignment occurs. A key issue is whether the sequence is known or predictable by the people involved in allocating participants to the comparison groups. The treatment allocation system should thus be set up so that the person enrolling participants does not know in advance which treatment the next person will receive, a process termed allocation concealment (item 18). Proper allocation concealment shields knowledge of forthcoming assignments, whereas proper random sequences (item 17) prevent correct anticipation of future assignments based on knowledge of past assignments (box 5).

#### **Box 5: Treatment allocation in trials**

The method used to assign interventions to trial participants is a crucial aspect of clinical trial design. Random assignment is the preferred method; it has been successfully used regularly in trials for more than 75 years.<sup>263</sup> Randomisation has three major advantages.<sup>264</sup> Firstly, when properly implemented, it eliminates selection bias, balancing both known and unknown prognostic factors, in the assignment of treatments. Without randomisation, treatment comparisons may be prejudiced, whether consciously or not, by selection of participants of a particular kind to receive a particular treatment. Secondly, random assignment permits the use of probability theory to express the likelihood that any difference in outcome between intervention groups reflects mere chance.<sup>265</sup> Thirdly, random allocation, in some situations, facilitates blinding the identity of treatments to the investigators, participants, and evaluators, possibly by use of a placebo, which reduces bias after assignment of treatments.<sup>266</sup> Of these three advantages, reducing selection bias at trial entry is usually the most important.<sup>267</sup>

Successful randomisation in practice depends on two inter-related aspects: adequate generation of an unpredictable allocation sequence and concealment of that sequence until assignment occurs.<sup>268 269</sup> A key issue is whether the sequence is known or predictable by the people involved in allocating participants to the comparison groups.<sup>270</sup> The treatment allocation system should thus be set up so that the person enrolling participants does not know in advance which treatment the next person will get, a process termed allocation concealment.<sup>268 269</sup> Proper allocation concealment shields knowledge of forthcoming assignments, whereas proper random sequences prevent correct anticipation of future assignments based on knowledge of past assignments.

Authors should provide sufficient information such that the reader can assess the methods used to generate the random allocation sequence and the likelihood of bias in group assignment. Any software used for random sequence generation should also be reported. It is important that information on the process of randomisation is included in the body of the main article and not as a separate supplementary file, where it can be missed by the reader.

The term “random” has a precise technical meaning. With random allocation, each participant has a known probability of receiving each intervention before one is assigned, and the assigned intervention is determined by a chance process and cannot be predicted. However, “random” is sometimes used inappropriately in the literature to describe trials in which non-random, “deterministic” allocation methods were used, such as alternation, hospital numbers, or date of birth. When investigators use such non-random methods, they should describe them precisely and should not use the term “random” or any variation of it. Even the term “quasi-random” is unacceptable for describing such trials. Trials based on non-random methods generally yield biased results<sup>4 268 272-277</sup>; bias presumably arises from the inability to adequately conceal these more predictable, non-random sequence generation systems.

Many methods of sequence generation are adequate. However, readers cannot judge adequacy from such terms as “random allocation,” “randomisation,” or “random” without further elaboration. Authors should specify the method of sequence generation, such as a random-number table or a computerised random number generator. The sequence may be generated by the process of minimisation, a non-random but generally acceptable method.

In some trials, participants are intentionally allocated in unequal numbers to each intervention: for example, to gain more experience with a new procedure or to limit costs of the trial. In such cases, authors should report the randomisation ratio (eg, 2:1, or two treatment participants per control participant; item 9).

In a representative sample of PubMed indexed trials in 2000, only 21% reported an adequate approach to random sequence generation<sup>220</sup>; this increased to 34% for a similar cohort of PubMed indexed trials in 2006.<sup>221</sup> Two more recent studies showed further small increases to about 40%,<sup>278 279</sup> but another reported a stubbornly similar level of 32%.<sup>277</sup> When authors report an adequate approach to random sequence generation, in over 90% of cases they report using a random number generator on a computer or a random number table.<sup>221 279</sup>

#### **Item 17b: Type of randomisation and details of any restriction (eg, stratification, blocking, and block size)**

*Examples*

“Treatment assignment was generated using a simple randomization scheme . . . given the open-label nature of the intervention to limit the potential bias due to predictable treatment assignment.”<sup>280</sup>

“Randomization (1:1) was performed by an independent researcher using computer generated random table numbers, with a block size of 20 and stratified for the indication of the IUI [intrauterine insemination] (mild male factor or unexplained subfertility).”<sup>281</sup>

“Participants were randomized at an individual-level (1:1 ratio) and were stratified by recruitment location (VU [Vrije University] and UvA [University of Amsterdam]). Block randomization was applied with randomly varied block sizes (6–12 allocations per block).”<sup>262</sup>

“Randomization was stratified by treatment centre, clinical severity (<4 vs >4 on a Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale standardized to range from 0 to 10), and by whether patients had previously received TENS [transcutaneous electrical nerve stimulation] with randomly varied block sizes of 2, 4, and 6.”<sup>282</sup>

“Randomization sequence was created using Stata 9.0 (StataCorp., College Station, TX) statistical software and was stratified by center with a 1:1 allocation using random block sizes of 2, 4, and 6.”<sup>283</sup>

#### *Explanation*

In trials of several hundred participants or more, simple randomisation can usually be trusted to generate similar numbers in the two trial groups<sup>284</sup> and to generate groups that are roughly comparable in terms of known and unknown prognostic variables.<sup>285</sup> For smaller trials of fewer than around 200 participants,<sup>286</sup> which are common, some form of restricted randomisation procedure to help achieve balance between groups in size or characteristics may be useful (box 6). However, larger trials of greater than approximately 200 participants may also benefit from registration. For example, they may stop before reaching their target size, they may need more power at interim analyses, or they may benefit from stratification with restriction.

It is important to indicate whether no restriction was used by stating such or by stating that simple randomisation was done. Otherwise, the methods used to restrict the randomisation, along with the method used for random selection, should be specified. For blocked randomisation, authors should provide details on how the blocks were generated (eg, by using a permuted block design with a computer random number generator), the block size or sizes, and whether the block size was fixed or randomly varied. If the trialists became aware of the block size(s), that information should also be reported as such knowledge could lead to them correctly deciphering future treatment assignments. Authors should specify whether stratification was used and, if so, which factors (eg, recruitment site, sex, disease stage) were involved; the categorisation cut-off thresholds within strata; and the method used for restriction. Although stratification is a useful technique, especially for smaller trials, it can be complicated to implement and may not perform as well as expected if many stratifying factors are used. If

**Box 6: Randomisation and minimisation**
**Simple randomisation**

Pure randomisation based on a single allocation ratio is known as simple randomisation. Simple randomisation with a 1:1 allocation ratio is analogous to a coin toss, although we do not advocate coin tossing for randomisation in a randomised trial. The term “simple” is somewhat of a misnomer. While other randomisation schemes sound complex and more sophisticated, in reality, simple randomisation is elegantly sophisticated in that it is more unpredictable and could surpass the bias prevention levels of all other alternatives.

**Restricted randomisation**

Restricted randomisation specifies any randomised approach that is not simple randomisation. Blocked randomisation is the most common form. Other means of restricted randomisation include replacement, biased coin, and urn randomisation, although these are used much less frequently.<sup>286</sup>

**Blocked randomisation**

Blocking can be used to ensure close balance of the numbers in each group at any time during the trial. After a block of every eight participants was assigned, for example, four would be allocated to each arm of the trial.<sup>287</sup> Improved balance comes at the cost of reducing the unpredictability of the sequence. Although the order of interventions varies randomly within each block, a person running the trial could deduce some of the next treatment allocations if they discovered the block size.<sup>288</sup> Blinding the interventions, using larger block sizes, and randomly varying the block size can ameliorate this problem.

**Stratified randomisation**

Stratification is used to ensure a good balance of participant characteristics in each group. By chance, particularly in small trials, trial groups may not be well matched for baseline characteristics, such as age and stage of disease. This weakens the trial’s credibility.<sup>289</sup> Such imbalances can be avoided without sacrificing the advantages of randomisation. Stratification ensures that the numbers of participants receiving each intervention are closely balanced within each stratum. Stratified randomisation is achieved by performing a separate randomisation procedure within each of two or more subsets of participants (eg, those defining each centre, age, or disease severity). Stratification by centre is common in multicentre trials. Stratification requires some form of restriction, such as blocking within strata. Stratification without some form of restriction is ineffective.

**Minimisation**

Minimisation improves balance between intervention groups for several selected patient factors (eg, age).<sup>271,290</sup> The first patient is truly randomly allocated; for each subsequent participant, the treatment allocation that minimises the imbalance on the selected factors between groups at that time is identified. That allocation may then be used, or a choice may be made at random with a heavy weighting in favour of the intervention that would minimise imbalance (eg, with a probability of 0.8). The use of a random component is generally preferable. Minimisation has the advantage of creating small groups closely similar in terms of measurable participant characteristics at all stages of the trial.

Minimisation offers the only acceptable alternative to randomisation, and some have argued that it is superior.<sup>291</sup> Conversely, minimisation lacks the theoretical basis for eliminating bias on all known and unknown factors. Nevertheless, in general, trials that use minimisation are considered methodologically equivalent to randomised trials, even when a random element is not incorporated.

minimisation (box 6) was used, it should be explicitly identified, as should the variables incorporated into the scheme; whether a random element was used should also be stated.

With blocking, although the order of interventions varies randomly within each block, individuals running the trial could deduce some of the future treatment allocations if they discovered the block size.<sup>288</sup> Discovering block sizes is much more likely in unblinded trials, where treatment allocations become known after assignment (box 6). Certain techniques, such as large block sizes and randomly varying block sizes, can help prevent the deciphering of future treatment allocations. Unfortunately, particularly with unblinded trials, a review “found that very few trials used techniques that would eliminate the risk of selection bias,” and that “These findings indicate that a substantial proportion of unblinded trials are at risk of selection bias.”<sup>292</sup> Indeed, in a recent study of 179 open, unblinded randomised trials, small block sizes were associated with subversion.<sup>293</sup>

Only 9% of 206 reports of trials in specialty journals<sup>269</sup> and 39% of 80 trials in general medical journals reported use of stratification.<sup>294</sup> In each case, only about half of the reports mentioned the use of restricted randomisation. Those studies and that of

Adetugbo and Williams<sup>295</sup> found that the sizes of the treatment groups in many trials were very often the same or quite similar, yet blocking or stratification had not been mentioned. One of a few possible causes of this close balance in numbers is under-reporting of the use of restricted randomisation, although non-random manipulation of treatment assignments is also suspected. A more recent study of 298 reports of trials in general medical journals found 69% reported the use of a stratified block method.<sup>296</sup>

**Item 18: Mechanism used to implement the random allocation sequence (eg, central computer/telephone; sequentially numbered, opaque, sealed containers), describing any steps to conceal the sequence until interventions were assigned**

*Examples*

“Participants were centrally assigned to randomised study treatment using an interactive web response system (IWRS) . . . Block randomisation schedules were computer generated by a vendor with a block size of 6 in a randomisation ratio of 2:1 and distributed to the IWRS vendor (endpointClinical) for participant randomisation.”<sup>297</sup>

“For allocation concealment, numbered containers were used. The interventions were sealed in

sequentially numbered identical opaque containers according to the allocation sequence.”<sup>298</sup>

“Furthermore, we employed syringes sequentially numbered and packaged in opaque and sealed containers. Specifically, syringes containing esmolol or placebo were centrally prepared, pre-coded based on the randomization list, and sent sequentially to the operating room immediately before administration.”<sup>299</sup>

“Allocation was concealed using sequentially numbered, opaque, sealed envelopes (SNOSE) prior to making the incision.”<sup>300</sup>

“Allocation concealment was done using sequentially numbered, sealed, opaque packages.”<sup>261</sup>

“The allocation sequence was concealed from the researcher (JR) enrolling and assessing participants in sequentially numbered, opaque, sealed and stapled envelopes. Aluminium foil inside the envelope was used to render the envelope impermeable to intense light. To prevent subversion of the allocation sequence, the name and date of birth of the participant was written on the envelope and a video tape made of the sealed envelope with participant details visible. Carbon paper inside the envelope transferred the information onto the allocation card inside the envelope and a second researcher (CC) later viewed video tapes to ensure envelopes were still sealed when participants’ names were written on them. Corresponding envelopes were opened only after the enrolled participants completed all baseline assessments and it was time to allocate the intervention.”<sup>301</sup>

#### *Explanation*

Item 17 discussed generation of an unpredictable sequence of assignments. Of considerable importance is how this sequence is applied when participants are enrolled into the trial (box 5). A generated allocation sequence should be implemented by using allocation concealment,<sup>269</sup> a critical mechanism that prevents foreknowledge of treatment assignment and thus shields those who enrol participants from being influenced by this knowledge. The decision to accept or reject a participant should be made, and informed consent should be obtained from the participant, in ignorance of the next assignment in the sequence.<sup>302</sup> In summary, adequate allocation concealment safeguards knowledge of forthcoming assignments, whereas proper random sequences (item 17) prevent correct anticipation of future assignments based on knowledge of past assignments.

Allocation concealment should not be confused with blinding (item 20). Allocation concealment seeks to prevent selection bias (box 5), protects the assignment sequence before and until allocation, and can always be successfully implemented.<sup>268</sup> In contrast, blinding seeks to prevent ascertainment bias, protects the sequence after allocation, and cannot always be implemented.<sup>269</sup> Without adequate allocation concealment, however, even random, unpredictable assignment sequences can be subverted.<sup>268 303</sup>

Centralised or third party assignment is especially desirable. Many good allocation concealment

mechanisms incorporate external involvement. Use of a pharmacy or central computer or telephone randomisation system are common techniques. Automated assignment systems are likely to become more common.<sup>221</sup> When external involvement is not feasible, an excellent method of allocation concealment is the use of numbered containers. The interventions (often medicines) are sealed in sequentially numbered identical containers according to the allocation sequence.<sup>304</sup> Enclosing assignments in sequentially numbered, opaque, sealed envelopes can be a good allocation concealment mechanism if it is developed and monitored diligently.<sup>288 305</sup> This method can be corrupted, however, particularly if it is poorly executed. Investigators should ensure that the envelopes are opaque when held to the light, and are opened sequentially and only after the participant’s name and other details are written on the appropriate sequentially numbered sealed envelope.<sup>288 305 306</sup>

A number of methodological studies provide empirical evidence to support these precautions.<sup>4268 275-277 307-309</sup> Trials in which the allocation sequence had been inadequately or unclearly concealed yielded larger estimates of treatment effects than did trials in which authors reported adequate allocation concealment. These findings provide strong empirical evidence that inadequate allocation concealment contributes to bias in estimating treatment effects.

Despite the importance of the mechanism of allocation concealment, published reports frequently omit such details. Among older studies, the mechanism used to allocate interventions was omitted in reports of 89% of trials on rheumatoid arthritis,<sup>310</sup> 48% of trials in obstetrics and gynaecology journals,<sup>269</sup> and 44% of trials in general medical journals.<sup>294</sup> In a more broadly representative sample of all PubMed indexed randomised trials, only 18% reported any allocation concealment mechanism and some of those reported mechanisms were inadequate.<sup>220</sup> At the same time, some trials where there is no reporting of allocation concealment may have been properly concealed, as demonstrated by inspection of their protocols.<sup>311</sup>

Newer studies further illuminate poor reporting of allocation concealment. Unclear reporting (ie, the authors did not provide sufficient information in the paper to allow judgment to be made on the adequacy of method of allocation concealment) was found in 78%<sup>312</sup> and 85%<sup>277</sup> of trials. Moreover, those two studies determined that only 27%<sup>312</sup> and 14%<sup>277</sup> used an adequate allocation concealment mechanism, while another found a similar level of 12%.<sup>278</sup> A review of trials in journals of traditional Chinese medicine found that only 7% used adequate allocation concealment.<sup>313</sup>

Fortunately, reporting and conduct may be improving in recent years, for example, after the CONSORT 2010 guidelines were published.<sup>314</sup> Another study found that reporting on allocation concealment and sequence generation was much better in journals that endorsed the CONSORT 2010 guidelines than in non-endorsing journals.<sup>305</sup> Moreover, that study found that 57% of trials in the sample had used an adequate

allocation concealment mechanism.<sup>305</sup> However, other empirical studies show only modest improvements, for example, an evaluation of over 176 000 trials<sup>315</sup> found that allocation concealment reporting increased from 5.1% in 1966-90 to 19.3% in 2010-18. While any improvement is encouraging, more efforts to improve conduct and reporting remain necessary.

**Item 19: Whether the personnel who enrolled and those who assigned participants to the interventions had access to the random allocation sequence**

*Examples*

“Sequential randomisation codes were computer generated using permuted blocks . . . An independent study statistician generated the randomisation codes . . . Good Clinical Practice (GCP) trained research nurses or First Nations health practitioners allocate the study medicine to each mother–infant pair by selecting the next sequentially labelled (prerandomised) study medication from the appropriate stratification group. The allocation sequence number is recorded by the research team on the data collection form (DCF), the database and in the participant’s medical record.”<sup>316</sup>

“LabCorp Drug Development (a subcontractor to the IWRS [interactive web response system] vendor) generated the live randomisation schedules. Site personnel enrolled participants in the IWRS. The IWRS assigned participants to the trial groups per live randomisation schedules. LabCorp Drug Development didn’t have any involvement in the rest of the trial . . . Site personnel were involved in participant care and performing trial procedures throughout the trial; however, they were masked to treatment assignment.”<sup>297</sup>

“All participants were screened by a masked trial doctor who also obtained informed consent . . . The masked trial doctor then provided all participants with guideline-recommended care and a prescription for the trial medicine kit . . . The randomisation sequence was created using randomly permuted blocks by an independent statistician (who had no involvement in the rest of the trial).”<sup>317</sup>

“The details of sequence generation and group allocation were unavailable to research team members bar one unblinded research assistant. The unblinded research team member that created the randomisation sequence had no contact with participants and was not involved with data collection or analysis.”<sup>318</sup>

“Block randomisation was by a computer generated random number list prepared by an investigator with no clinical involvement in the trial . . . After the research nurse had obtained the patient’s consent, she telephoned a contact who was independent of the recruitment process for allocation consignment.”<sup>319</sup>

*Explanation*

As noted in item 18, concealment of the allocated intervention at the time of enrolment is especially important. Thus, in addition to knowing the methods used, it is also important to understand how the random sequence was implemented; specifically, whether the personnel who enrolled and those who assigned participants to the interventions had access to the random allocation sequence.

In practice, the process of randomising participants into a trial has three different steps: sequence generation, allocation concealment mechanism, and implementation (table 6). Although the same individuals may carry out more than one process under each heading, investigators should strive for complete separation of the people involved with generation and allocation concealment from the people who implement assignments. Thus, if someone is involved in the sequence generation or allocation concealment steps, ideally, they should not be involved in the implementation step. When this separation is not possible, importantly the investigators should ensure that the assignment schedule is unpredictable and locked away from even the person who generated it.

Even with flawless sequence generation and allocation concealment, failure to separate creation and concealment of the allocation sequence from assignment to trial group may introduce bias. For example, the person who was enrolling and assigning participants and who also generated an allocation sequence would likely have access to the sequence list and could consult it when interviewing potential participants for a trial. Thus, that person could bias the enrolment or assignment process, regardless of the unpredictability of the assignment sequence and the impenetrability of the allocation concealment mechanism. Investigators must therefore ensure that the assignment schedule is unpredictable and locked away (eg, in a safe deposit box in a building inaccessible to the enrolment location) from even the person who generated it. In that instance, the report of the trial should specify where the investigators stored the allocation list.

Thus, for full assessment of randomisation in a trial report, authors should confirm that personnel who enrolled and those who assigned participants to the interventions did not have access to the random allocation sequence. At minimum, authors should confirm complete separation of the people involved with generation and allocation concealment from the people involved in the implementation of assignments. If complete separation did not occur, then authors should describe how the people involved in the implementation were prevented from accessing the

**Table 6 | Steps in a typical randomisation process**

Sequence generation	Allocation concealment	Implementation
Generate allocation sequence by a random procedure	Develop allocation concealment mechanism (eg, numbered, identical bottles or sequentially numbered, sealed, opaque envelopes)	Enrol participants: <ul style="list-style-type: none"> <li>assess eligibility</li> <li>discuss the trial</li> <li>obtain informed consent</li> <li>enrol participant in trial</li> </ul>
	Prepare the allocation concealment mechanism using the allocation sequence from the sequence generation step	Ascertain intervention assignment (eg, opening next envelope) and assign participant to appropriate intervention Administer intervention

sequence (eg, specifying that the allocation sequence was locked in a secure location).

Sometimes those who enrol and those who assign are different people, but frequently the same individuals do both. These individuals may be termed differently by authors (eg, recruiters), but the functions remain the same.

Full assessment of implementation is not always possible from trial reports. In a sample of 199 medical journals, 63% of trial reports did not provide sufficient information to assess whether the person who generated the allocation sequence was not also the person who allocated participants to treatment groups.<sup>306</sup> Only 31% of trial reports provided sufficient details on who recruited participants and who generated the allocation sequence.<sup>306</sup>

**Item 20a: Who was blinded after assignment to interventions (eg, participants, care providers, outcome assessors, data analysts)**

*Examples*

“Whereas patients and physicians allocated to the intervention group were aware of the allocated arm, outcome assessors and data analysts were kept blinded to the allocation.”<sup>320</sup>

“Blinding and equipoise were strictly maintained by emphasizing to intervention staff and participants that each diet adheres to healthy principles, and each is advocated by certain experts to be superior for long-term weight-loss. Except for the interventionists (dieticians and behavioural psychologists), investigators and staff were kept blind to diet assignment of the participants.

The trial adhered to established procedures to maintain separation between staff that take outcome measurements and staff that deliver the intervention. Staff members who obtained outcome measurements were not informed of the diet group assignment. Intervention staff, dieticians and behavioural psychologists who delivered the intervention did not take outcome measurements. All investigators, staff, and participants were kept masked to outcome measurements and trial results.”<sup>321</sup>

“This was a double-blind study with limited access to the randomisation code . . . The treatment each patient received was not disclosed to the investigator, study site staff, patient, sponsor personnel involved with the conduct of the study (with the exception of the clinical supply staff and designated safety staff), or study vendors.”<sup>297</sup>

“Physicians, patients, nurses responsible for referring the patients, the statistician, also the investigators who rated the patients and administered the drugs, were all blinded to the allocation.”<sup>261</sup>

*Explanation*

The term “blinding” (masking) refers to withholding information about the assigned interventions from people involved in the trial who may potentially be influenced by this knowledge. Blinding is an important safeguard against bias, particularly when assessing subjective outcomes.<sup>308</sup>

Benjamin Franklin has been credited as being the first to use blinding in a scientific experiment.<sup>322</sup> He blindfolded participants so they would not know when he was applying mesmerism (a popular healing technique of the 18th century) and in so doing demonstrated that mesmerism was a sham. Since then, the scientific community has widely recognised the power of blinding to reduce bias, and it has remained a commonly used strategy in scientific experiments.

Box 7 on blinding terminology defines the groups of individuals (ie, participants, healthcare providers, data collectors, outcome assessors, and data analysts) that can potentially introduce bias into a trial through knowledge of the treatment assignments. Participants may respond differently if they are aware of their treatment assignment (eg, respond more favourably when they receive the new treatment).<sup>308</sup> Lack of blinding may also influence adherence with the intervention, use of co-interventions, and risk of dropping out of the trial.

Unblinded healthcare providers may introduce similar biases; and unblinded data collectors may differentially assess outcomes (eg, frequency or timing), repeat measurements of abnormal findings, or provide encouragement during performance testing. Unblinded outcome assessors may differentially assess subjective outcomes, and unblinded data analysts may introduce bias through the choice of analytical strategies, such as the selection of favourable time points or outcomes and by decisions to remove patients from the analyses. These biases have been well documented.<sup>32 308 325 327-329</sup>

#### Box 7: Blinding terminology

For a technical term to be useful, its use and interpretation must be consistent. Authors of trials commonly use the term “double blind,” and less commonly the terms “single blind” or “triple blind.” A problem with this lexicon is that there is great variability in clinician interpretations and epidemiological textbook definitions of these terms.<sup>323</sup> Moreover, a study of 200 randomised trials reported as double blind demonstrated 18 different combinations of groups actually blinded when the authors of these trials were surveyed, and approximately one in every five of these trials—reported as double blind—did not blind participants, healthcare providers, or data collectors.<sup>324</sup>

This research demonstrates that terms are ambiguous and, as such, authors and editors should abandon their usage in isolation without defining them. Authors should instead explicitly report the blinding status of the people involved for whom blinding may influence the validity of a trial.

The healthcare providers include all personnel (eg, physicians, chiropractors, physiotherapists, nurses) who care for the participants during the trial. Data collectors are the individuals who collect data on the trial outcomes. Outcome assessors are the individuals who determine whether a participant did experience the outcomes of interest.

Some researchers have also advocated blinding and reporting the blinding status of the data monitoring committee and the manuscript writers.<sup>325</sup> Blinding of these groups is uncommon and the value of blinding them is debated.<sup>326</sup>

Sometimes one group of individuals (eg, the healthcare providers) is also the same individuals fulfilling another role in a trial (eg, the data collectors). Even if this is the case, the authors should state the blinding status of these groups to allow readers to judge the validity of the trial.

Blinding, unlike allocation concealment (item 18), may not always be appropriate or possible. In pragmatic trials (trials that try to make the experience as close as real life so as to understand real world effectiveness), blinding of participants and healthcare providers would decrease the pragmatism of the trials, since patients in real life are not blinded.<sup>330</sup> An example where blinding is impossible is a trial comparing levels of pain associated with sampling blood from the ear or thumb.<sup>331</sup> However, in randomised trials for which blinding is possible, lack of blinding has usually been associated with empirical evidence of exaggeration in treatment effect estimates.<sup>276 308 332-336</sup> Blinding is particularly important when outcome measures involve some subjectivity, such as assessment of pain. Yet, blinding may not be as important in certain fields or with certain outcomes. For example, blinding of data collectors and outcome assessors is unlikely to matter for objective outcomes, such as death from any cause. Indeed, some methodological investigations have not found that lack of blinding is associated with empirical evidence of bias in treatment effect estimates.<sup>337-343</sup> Even then, however, lack of participant or healthcare provider blinding can lead to other problems, such as differential attrition.<sup>344</sup> In certain trials, especially surgical trials, blinding of participants and healthcare providers is often difficult or impossible, but blinding of data collectors and outcome assessors for both benefits and harms is often achievable and recommended. For example, lesions can be photographed before and after treatment and assessed by an external observer.<sup>345</sup> Regardless of whether blinding is possible, authors can and should always state who was blinded (ie, participants, healthcare providers, data collectors, data analysts, and/or outcome assessors).<sup>346 347</sup>

However, authors frequently do not report whether blinding was used.<sup>348 349</sup> For example, reports of 51% of 506 trials in cystic fibrosis,<sup>350</sup> 33% of 196 trials in rheumatoid arthritis,<sup>310</sup> and 38% of 68 trials in dermatology<sup>295</sup> did not state whether blinding was used. Similarly, a more recent review found that the reports of 38% of 622 trials in high impact anaesthesiology journals did not explicitly describe the trial as blinded or non-blinded.<sup>351</sup> Moreover, when describing some form of blinding, the most used term was the ambiguous “double blind.”<sup>351</sup> Authors should explicitly state who was blinded, but only 14% of 622 trials explicitly reported whether the three key groups of individuals—that is, the participants, healthcare providers, and data collectors—were blinded or not.<sup>351</sup> The rate did improve from 10% to 26% over the years of that review, but more improvement is needed. Until authors of trials improve their reporting of blinding, readers will have difficulty in judging its adequacy.

The term “masking” is sometimes used in preference to “blinding” to avoid confusion with the medical condition of being without sight. However, “blinding” in its methodological sense appears to be more universally understood worldwide and to be generally preferred for reporting clinical trials.<sup>344 345 352</sup>

**Item 20b: If blinded, how blinding was achieved and description of the similarity of interventions Examples**

“Jamieson Laboratories Inc. provided 500-mg immediate release niacin in a white, oblong, bisect caplet. We independently confirmed caplet content using high performance liquid chromatography . . . The placebo was matched to the study drug for taste, color, and size, and contained microcrystalline cellulose, silicon dioxide, dicalcium phosphate, magnesium stearate, and stearic acid.”<sup>353</sup>

“Placebo tablets were identical to NAC [N-acetylcysteine] tablets in color, shape, size, and odor. They were all kept in identical containers and were administered by an investigational drug pharmacist.”<sup>261</sup>

“The study treatment and placebo tablets and bottles were identical in physical appearance . . . The IWRs [interactive web response system] housed treatment codes and bottle numbers for study treatment. In case of an emergency, the investigator had the sole responsibility for determining if unmasking of a participant’s treatment assignment was warranted to provide appropriate medical care. Participant safety was always the first consideration in making such a determination. The IWRs was programmed with blind-breaking instructions to guide the investigator on how to obtain treatment assignment in the event of an emergency unmasking. The investigator was requested to contact the medical monitor promptly in case of any treatment unmasking. If a participant’s treatment assignment was unmasked, the sponsor was to be notified within 24 h after unmasking. The date and reason for the unmasking were recorded in the source documentation and electronic case report form, as applicable. Investigators broke the masking for four participants: one in ELEVATE UC 12 (on etrasimod) and three in ELEVATE UC 52 (on etrasimod).”<sup>297</sup>

*Explanation*

Blinding of participants, healthcare providers, data collectors, and outcome assessors in a trial requires adequate procedures to both achieve and maintain blinding.<sup>346 347</sup> Just as we seek evidence of adequate allocation concealment to assure us that assignment was truly random, we seek evidence on the method of blinding.

If researchers contend that the trial investigators, participants, and assessors were blinded, then they should provide information about the mechanism used to establish blinding (eg, placebo identical to the experimental intervention, sham intervention, sham surgery).<sup>346 347</sup> They should describe the similarity of treatment characteristics (eg, route of administration, appearance, smell, taste) and where relevant methods used to mask some characteristics of the treatments (eg, use of special flavours to mask a distinctive taste, opaque coverage to conceal intravenous treatments with different appearances, double-dummy procedures).<sup>346 347</sup>

Blinding can be difficult to maintain over time because of dosage adaptation over time or the occurrence of specific side effects. Specific procedures to maintain blinding can be implemented (eg, centralised assessment of side effects, centralised adapted dosage, or provision of sham results of complementary investigations).

Even if blinding of participants and healthcare providers is not possible, blinding data collectors and outcome assessors could still be implemented to limit ascertainment bias. This could be achieved, for example, through centralised assessment of complementary investigation (eg, anonymised radiography), physician mediated data (eg, video, photography, audiotape), and clinical events (eg, adjudication of clinical events from extract of the case report form).

Details of how blinding was achieved are important because slight, but discernible, differences between interventions can lead to large problems in bias. Notably, inadequate matching related to discernible differences in colour and taste seem particularly problematic.<sup>354</sup> It is important that authors report any known compromises in blinding. For example, authors should report if it was necessary to unblind any participants at any point during the conduct of the trial. Moreover, authors should report the risk of unblinding but, unfortunately, such reporting is rare. In a random sample of 300 publications describing blinded randomised clinical trials indexed in PubMed, only 8% reported on risk of unblinding.<sup>355</sup> It is also important to report any procedures, pretrial or concurrent, that are intended to reduce or evaluate risk of compromised blinding.<sup>354-356</sup> Indeed, some pretrial assessments of unblinding may be helpful in reducing the risk of unblinding in the eventual randomised trial. Thus, authors of randomised trial articles should report procedures to avoid, document, and address cases of overt unblinding.<sup>354-355</sup>

Where appropriate, authors should also describe any procedures used for emergency unblinding (ie, disclosing the assigned intervention of a trial participant for specific reasons such as harms). They should indicate whether they used fixed code to indicate group assignment (eg, A=group 1; B=group 2) or a unique code for each participant. Use of a fixed code will increase the risk of unblinding because unblinding a participant could result in unblinding several or all trial participants.<sup>6,286,357</sup>

Some people have advocated testing for blinding by asking participants or healthcare providers at the end of a trial whether they think the participant received the experimental or control intervention.<sup>358</sup> Because participants and healthcare providers will frequently know whether the participant has experienced the primary outcome, this makes it difficult to determine whether their responses reflect failure of blinding or accurate assumptions about the efficacy of the intervention.<sup>359</sup> Thus, given the uncertainty this type of information provides, the usefulness of tests of blinding has been questioned.<sup>354,355,360</sup> Testing for

blinding was not an included item in CONSORT 2010, and still is not in CONSORT 2025. Nevertheless, if investigators decide to conduct tests of blinding, we encourage them to completely report their findings with appropriate limitations.<sup>356</sup>

#### **Item 21a: Statistical methods used to compare groups for primary and secondary outcomes, including harms**

##### *Examples*

"The primary outcome was analysed using a mixed effects log-binomial model to generate an adjusted risk ratio (RR) and an adjusted risk difference (using an identity link function), including centre as a random effect. Statistical significance of the treatment group parameter was determined (p value generated) through examination of the associated  $\chi^2$  statistic (obtained from the log-binomial model which produced the RR). Binary secondary outcomes were analysed as per the primary outcome. Time to hCG [human chorionic gonadotrophin] resolution was considered in a competing risk framework to account for participants who had surgical intervention for their ectopic pregnancy. A cumulative incidence function was used to estimate the probability of occurrence (hCG resolution) over time. A Fine and Gray model was then used to estimate a subdistribution adjusted hazard ratio (HR) directly from the cumulative incidence function. In addition, a further Cox proportional hazard model was fitted and applied to the cause-specific (non-surgical resolution) hazard function and used to generate an adjusted HR. Return to menses was analysed using a Cox regression model. Number of hospital visits associated with treatment was analysed using a Poisson regression model, including centre as a random effect to generate an adjusted incidence ratio."<sup>255</sup>

"For the primary continuous outcome and secondary outcomes, linear mixed-effect models were used, with outcome measurement (at the two follow-up timepoints) as the dependent variable. The models included fixed effects for timepoint, treatment, timepoint by treatment interactions, the baseline measure of the outcome, and therapist, assuming a linear relationship between baseline and outcome. The dichotomous outcome of recovery in the delusion was analysed using a logistic mixed-effect model. Persecutory delusion conviction was analysed as a continuous and also as a dichotomous (recovery) variable. The models included a random intercept for participant, an unstructured correlation matrix for the residuals, and were fitted using restricted maximum likelihood estimation . . . For each outcome and timepoint, we report the treatment effect estimate as the adjusted mean difference between groups, its SE [standard error], 95% CIs [confidence intervals], and p value. In addition, we report estimates for Cohen's d effect sizes as the adjusted mean difference of the outcome (between the groups) divided by the sample SD [standard deviation] of the outcome at baseline."<sup>361</sup>

“Analyses followed a prespecified statistical analysis plan. The primary outcome (ODQ [Oswestry Disability Questionnaire] score at 18 weeks after randomisation) was compared between groups with a linear regression model, adjusted for baseline ODQ, with centre as a random effect. ODQ score, visual analogue scores (VAS) for back pain, VAS for leg pain, MRM [modified Roland-Morris] outcome score, and COMI [Core Outcome Measures Index] score at all follow-up visits were analysed with a repeated measures mixed-effects model adjusting for baseline outcome measure, treatment group, time (as a continuous variable), and a time-treatment arm interaction (if significant). Centre and participant were random effects in the repeated measures models. A second model adjusted for other prespecified variables, age, sex, duration of symptoms, body-mass index, and size of disc prolapse (as a percentage of the diameter of the spinal canal, categorised as <25%, 25–50%, or >50%).”<sup>362</sup>

“We analysed the primary outcome (between-group difference in the SPPB [short physical performance battery] at 12 months) using linear mixed models, adjusted for baseline measurements, minimisation variables (age, sex and CKD [chronic kidney disease] category) and a random effect variable for recruitment site. We analysed secondary outcomes using repeated measures mixed models, including all participants and including data from all available timepoints. Models were adjusted for baseline values and the minimisation variables. We conducted time-to-event analyses (time to death, time to commencing renal replacement therapy) using Cox proportional hazards models adjusted for minimisation variables. All participants were included in these analyses, with participants censored at the point of dropout or truncation of follow-up for those not reaching the analysis endpoint before 24 months. For all analyses, we took a two-sided p value of <0.05 as significant with no adjustment for multiple testing.”<sup>363</sup>

#### *Explanation*

Various methods can be used to analyse data, and it is crucial to ensure that the chosen approach is suitable for the specific context. Specifying the statistical procedures and software used for each analysis is essential, and additional clarification may be required in the results section of the report. Authors should describe the statistical methods insufficient detail to allow a knowledgeable reader with access to the original data to verify the reported results, as emphasised by the ICMJE (<https://www.icmje.org/>). It is also important to elaborate on specific aspects of the statistical analysis, such as the intention-to-treat approach.

Details of all statistical analyses are frequently prespecified in a statistical analysis plan, a document that accompanies the trial protocol. In the report of the trial results, authors should detail and justify any deviation from the statistical analysis plan or from the protocol if no statistical analysis plan was developed.

They should clarify which analyses were prespecified and which were post hoc.

Most analysis approaches provide an estimate of the treatment effect, representing the difference in outcomes between comparison groups, and authors should also indicate the effect measure (eg, absolute risk) considered. Authors should accompany this with a CI for the estimated effect, delineating a central range of uncertainty regarding the actual treatment effect. The CI may be interpreted as the range of values for the treatment effect that is compatible with the observed data. Typically, a 95% CI is presented, signifying the range anticipated to encompass the true value in 95 of 100 similar studies.

Study findings can also be assessed in terms of their statistical significance. The P value represents the probability that the observed data (or a more extreme result) could have arisen by chance when the interventions did not truly differ. The statistical significance level that will be used should be reported. In the results section, actual P values (for example, P=0.001) are strongly preferable to imprecise threshold reports such as P<0.05.<sup>364 365</sup>

Some trials may use bayesian methods.<sup>366-369</sup> In this case, the choices of priors, computational decisions, and any modelling methods used should be described. Most bayesian trials so far have been for early phases of drug development, but this approach can be applicable to any phase. Typically, results are presented as treatment effects along with credible intervals.

Where an analysis lacks statistical power (eg, harms outcomes), authors may prefer descriptive approaches over formal statistical analysis.

While the necessity for covariate adjustments is generally reduced in randomised trials compared with epidemiological studies, considering an adjusted analysis can have value in terms of increased power and precision, particularly if there is an indication that one or more variables may have prognostic value.<sup>370</sup> It is preferable for adjusted analyses to be explicitly outlined in the study protocol (item 3). For instance, it is often advisable to make adjustments for stratification variables,<sup>371</sup> in keeping with the principle that the analysis strategy should align with the study design. In the context of randomised trials, the decision to make adjustments should not be based on whether there are baseline covariates that are statistically significantly different between randomised groups. The testing of baseline imbalance in covariates should be avoided,<sup>370</sup> as if randomisation is properly conducted, then by definition, any differences in baseline covariates between treatment arms are random. The rationale for any adjusted analyses and the statistical methods used should be specified, along with clarifying the choice of covariates that were adjusted for, indicating how continuous variables were handled (eg, linear, modelled with splines),<sup>372</sup> and specifying whether the analysis was planned or post hoc. Reviews of published studies show that reporting of adjusted analyses is inadequate with regard to all of these aspects.<sup>373 374</sup>

Multiplicity issues are prevalent in trials and merit special consideration, especially in cases involving multiple primary outcomes, multiple time points stemming from repeated assessments of an outcome, multiple planned analyses for an outcome (such as interim or subgroup analyses (item 21d)), or analyses of numerous secondary outcomes (see CONSORT outcomes extension for more details).<sup>21</sup> Any methods used to mitigate or account for multiplicity should be described. If no methods have been used to account for multiplicity (eg, not applicable, or not considered), then this should also be reported, particularly when a large number of analyses has been carried out.

**Item 21b: Definition of who is included in each analysis (eg, all randomised participants), and in which group**

*Examples*

"The primary statistical analyses were performed according to the treatment to which the participants were randomly assigned. The analyses of the efficacy and safety outcomes (other than adverse events) included all available data from all randomized participants who contributed at least 1 value after baseline for the outcome of interest. The data that were obtained after a participant enrolled in another trial of an investigational treatment were excluded from the analyses. However, the participant was included in the analyses if that participant contributed at least 1 value after baseline for the outcome of interest prior to enrolling in another trial."<sup>375</sup>

"Efficacy outcomes were assessed using intention-to-treat analysis (ie, the full set of all randomly assigned patients). Safety outcomes were assessed using the safety analysis set of all randomly allocated patients exposed to at least one dose of randomised intervention."<sup>376</sup>

"Efficacy analyses and other exploratory analyses were performed in the full analysis set (defined as all patients randomly assigned to the study, including those who did not receive a dose of study treatment). Safety analyses were performed in the safety analysis set (defined as patients who received at least one dose of study treatment). The per protocol set was defined as all patients in the full analysis set who complied with the protocol in terms of exposure to study treatment, availability of tumour assessments, and absence of major protocol deviations likely to affect efficacy outcomes. Sensitivity analyses of the primary endpoint were performed on the per protocol analysis set."<sup>377</sup>

"The primary analysis population was defined as all participants who completed baseline and 36-week assessments. The primary analysis of the primary outcome, AMCA [amended motor club assessment] score at 36 weeks, followed a modified intention-to-treat approach, regardless of compliance to the intervention, but did exclude patients who were deemed ineligible after randomisation, those who withdrew from the trial and were unwilling for their previously collected data to be used, or those who did not provide baseline and week 36 measurements."<sup>378</sup>

*Explanation*

A key strength of a randomised trial design is the avoidance of bias when randomly allocating trial participants to interventions. To preserve the benefits of randomisation, all randomised participants are included in the analysis and retained in the group to which they were allocated. Meeting these two conditions defines an intention-to-treat analysis—which is widely recommended as the preferred analysis strategy.<sup>379-381</sup> However, strict adherence to an intention-to-treat analysis is often difficult to achieve owing to missing outcomes for some trial participants (item 21c) or non-adherence to the trial intervention protocol. While imputation of missing outcomes would allow an intention-to-treat analysis, it does not guarantee an avoidance of bias except under strong assumptions about the missing data which may be unknown.

Various strategies for performing intention-to-treat analyses in the presence of missing outcome data are available.<sup>382</sup> When the number of missing outcomes is not large, the analysis population could be all randomised participants with outcome observed (known as an "available case" population) under a plausible missing data mechanism, and sensitivity analyses could be performed exploring departures from this assumption (thereby using all randomised participants at least in sensitivity analyses).<sup>382</sup> Concerns may arise when the frequency or the causes of dropping out differ between the intervention groups. Striving for intention-to-treat analysis by imputing values for participants with missing outcomes may lead to use of inadequate methods such as last observation carried forward.<sup>382-385</sup>

Regardless of whether all randomised participants (completely observed outcomes or imputed outcomes) or a subset of randomised participants with observed outcomes are included in the primary analysis, the analysis population should be described. Authors often describe performing analyses on a "modified intention-to-treat" population to cover departures from a strict intention-to-treat that excludes participants who did not adequately adhere to the protocol such that they did not receive some minimum amount of the intervention—in such cases, what defines the minimum amount of the intervention should be explained (eg, those participants receiving at least one dose of the medication). It is also common to include analyses based on a per protocol population, which includes participants completing the study with no major protocol deviations. Excluding participants may compromise the randomisation and lead to biased estimates of treatment effects if appropriate methods are not used. Other analysis populations are possible (eg, a safety population), and their rationale and definition should be explained. Thus, authors should clearly define which participants are included in each analysis and in which intervention group and avoid terms such as "modified intention-to-treat" or "per protocol" analysis.

**Item 21c: How missing data were handled in the analysis**

*Examples*

“Regarding the multiple imputation procedure, briefly, for each outcome, the analysis model used was a linear regression with treatment arm, baseline outcome, and ethnicity (randomization stratifier) as explanatory variables. The imputation models contained all the variables of the analysis model(s) as well as factors associated with missingness: age (identified empirically to predict missingness,  $P=.03$ ) and adherence (number of doses taken of either vitamin D or placebo,  $P<.001$ ).”<sup>386</sup>

“To consider the potential impact of missing data on trial conclusions, we used multiple imputation (data missing at random) and sensitivity analysis (data not missing at random). Multiple imputation by chained equations was performed using the “mi impute chained” command in Stata. We used a linear regression model to impute missing outcomes for the HOS ADL [activities of daily living subscale of the hip outcome score] at eight months post-randomisation. Variables in the imputation model included all covariates in the analysis model (baseline HOS ADL (continuous), age (continuous), and sex). In addition, we included other variables that were thought to be predictive of the outcome (lateral centre-edge angle, maximum  $\alpha$  angle, Kellgren-Lawrence grade, and baseline HADS score). Imputations were run separately by treatment arm and based on a predictive mean matching approach, choosing at random one of the five HOS ADL values with the closest predicted scores. Missing data in the covariates that were included in the multiple imputation model were imputed simultaneously (multiple imputation by chained equation approach). Sensitivity analysis was performed using the “rctmiss” command in Stata, and we considered scenarios where participants with missing data in each arm were assumed to have outcomes that were up to 9 points worse than when data were missing at random.”<sup>242</sup>

“Analyses for the 2 primary outcomes compared each treatment with usual care using multiple imputation to handle missing data and a Bonferroni-corrected 2-tailed type I error of .025. We performed 20 imputations with a fully conditional specification using Proc MI in SAS. Imputation was performed with the following prespecified variables: age, study group, study site, clinic, sex, race and ethnicity, body mass index, exercise frequency at baseline, education, employment status, smoking status, other medical conditions at baseline, number of medications used for spine pain at baseline, duration of pain at baseline, number of previous pain episodes, STarT Back score, baseline ODI, baseline self-efficacy, baseline EQ-5D-5L, and scores for patient-reported outcomes at every follow-up point (ODI [Oswestry Disability Index], cost, Lorig et al self-efficacy scale, and EQ-5D-5L [EuroQoL 5-dimensional 5-level questionnaire]). Each imputed data set was analyzed separately using Proc GENMOD in SAS (with an identity link and normally distributed

errors for ODI and a log link and Poisson-distributed errors for spine-related spending).”<sup>387</sup>

“Missing peak  $\dot{V}O_2$  [oxygen consumption] data at week 20, regardless of the type of intercurrent event, was imputed using multiple imputation methodology under the missing at random assumption for the primary analysis. Sensitivity analyses were performed by exploring a missing not at random assumption in the imputation of peak  $\dot{V}O_2$ . The imputation model used a regression multiple imputation, which includes treatment group, baseline respiratory exchange ratio, persistent atrial fibrillation (yes or no), age, sex, baseline peak  $\dot{V}O_2$ , baseline hemoglobin level, baseline estimated glomerular filtration rate, baseline body weight, baseline KCCQ [Kansas City cardiomyopathy questionnaire] total symptom score, baseline NYHA [New York Heart Association] class, and baseline average daily activity units (refers to 10 hours of wearing during the awake time for  $\geq 7$  days unless otherwise specified). Treatment group, persistent atrial fibrillation (yes or no), baseline NYHA class, and sex were treated as categorical variables. Fifty imputed data sets were generated. Each of the imputed data sets was analyzed using the analysis of covariance model of the primary analysis. Least square mean (LSM) treatment difference and the standard error were combined using Rubin’s rules to produce an LSM estimate of the treatment difference, its 95% CI [confidence interval], and P value for the test of null hypothesis of no treatment effect.”<sup>388</sup>

“Multiple imputation was preplanned for the primary outcome measure in the case of missing data; however, because there were no missing data relating to ventilator-free days, imputation was not required.”<sup>389</sup>

*Explanation*

Missing data are common when conducting medical research. Collecting data on all study participants can be challenging even in a trial that has mechanisms to maximise data capture. Missing values can occur in either the outcome or in one or more covariates, or usually both. There are many reasons why missing values occur in the outcome. Patients may stop participating in the trial, withdraw consent for further data collection, or fail to attend follow-up visits; all of which could be related to the treatment allocation, specific (prognostic) factors, or experiencing a specific health outcome.<sup>390</sup> Missing values could also occur in baseline variables, such that all the necessary data needed to conduct the trial have been only partially recorded. Despite the ubiquity of missing data in medical research, the reporting of missing data and how they are handled in the analyses is poor.<sup>391-399</sup>

Many trialists exclude patients without an observed outcome. Once any randomised participants are excluded, the analysis is not strictly an intention-to-treat analysis. Most randomised trials have some missing observations. Trialists effectively must choose between omitting the participants without final outcome data, imputing their missing outcome data, or using model based approaches such as fitting a linear

mixed model to repeated measures data.<sup>371</sup> A complete case (or available case) analysis includes only those participants whose outcome is known. While a few missing outcomes will not cause a problem, many trials have more than 10% of randomised patients with missing outcomes.<sup>391-398</sup> This common situation will result in loss of power by reducing the sample size, and bias may well be introduced if being lost to follow-up is related to a participant's response to treatment. There should be concern when the frequency or the causes of dropping out differ between the intervention groups.

Participants with missing outcomes can be included in the analysis if their outcomes are imputed (ie, their outcomes are estimated from other information that was collected) or if using a model based approach. Imputing the values of missing data allows the analysis to potentially conform to intention-to-treat analysis but requires strong assumptions, which may be hard to justify. Simple imputation methods are appealing, but their use may be inadvisable as they fail to account for uncertainty introduced by missing data and may lead to invalid inferences (eg, estimated standard errors for the treatment effect will be too small).<sup>400</sup> For randomised trials with missing data within repeated measures data, model based approaches such as fitting a linear mixed model can be used to estimate the treatment effect at the final time point which is valid under a missing-at-random assumption. A model is fit at a (limited) number of time points following

randomisation, by including fixed effects for time and randomised group and their interaction.<sup>371</sup>

Another approach that is sometimes used is known as "last observation carried forward," in which missing final values of the outcome variable are replaced by the last known value before the participant was lost to follow-up. Although this method might appear appealing through its simplicity, the underlying assumption will rarely be valid, so the method may introduce bias, and makes no allowance for the uncertainty of imputation. The approach of last observation carried forward has been severely criticised.<sup>383 384 401</sup> Sensitivity analyses should be reported to understand the extent to which the results of the trial depend on the missing data assumptions and subsequent analysis (item 21d).<sup>402</sup> When the findings from the sensitivity analyses are consistent with the results from the primary analysis (eg, complete case for the primary analysis and multiple imputation for a sensitivity analysis), trialists can be reassured that the missing data assumptions and associated methods had little impact on the trial results.<sup>403</sup>

Regardless of what data are missing, how such data are to be analysed and reported needs to be carefully planned. Authors should provide a description on how missing data were handled in sufficient detail to allow for the analysis to be reproduced (in principle; box 8).

#### **Box 8: Guidance for reporting analytical approaches to handle missing data (adapted from Hussain et al<sup>404</sup>)**

##### **Methods**

- Report any strategies used to reduce missing data throughout the trial process.
- Report if and/or how the original sample size calculation accounted for missing data (item 16a) and the justification for these decisions. Report if and/or how the sample size was reassessed during the course of the trial (item 16b).
- Report the assumption about the missing data mechanism for the primary analysis and the justification for this choice, for all outcomes. For multiple imputation methods, report<sup>405</sup>:
  - What variables were included in the imputation procedure?
  - How were non-normally distributed and binary/categorical variables dealt with?
  - If statistical interactions were included in the final analyses (item 21a), were they also included in imputation models?
  - Was imputation done separately or by randomised group?
  - How many imputed datasets were created?
  - How were results from different imputed datasets combined?
- Report the method used to handle missing data for the primary analysis (eg, complete case, multiple imputation) and the justification for the methods chosen, for all outcomes. Include whether or which auxiliary variables were collected and used.
- Report the assumptions about the missing data mechanism (eg, missing at random) and methods used to conduct the missing data sensitivity analyses for all outcomes, and the justification for the assumptions and methods chosen.
- Report how data that were truncated due to death or other causes were handled with a justification for the method(s) (if relevant).

##### **Results**

- Report the numbers and proportions of missing data in each trial arm.
- Report the reasons for missing data in each trial arm.
- Report a comparison of the characteristics of those with observed and missing data.
- Report the primary analysis based on the primary assumption about the missing data mechanism, for all outcomes.
- Report results of the missing data sensitivity analyses for all outcomes. As a minimum, a summary of the missing data sensitivity analyses should be reported in the main paper with the full results in the supplementary material.

##### **Discussion**

- Discuss the impact of missing data on the interpretation of findings, considering both internal and external validity. For multiple imputation, include whether the variables included in the imputation model make the missing-at-random assumption plausible.

**Item 21d: Methods for any additional analyses (eg, subgroup and sensitivity analyses), distinguishing prespecified from post hoc**

*Examples*

"We conducted prespecified sensitivity analyses to examine the effect of our assumption that participants who withdrew or were lost to follow-up returned to smoking: (1) a complete case analysis and (2) multiple imputation to impute missing smoking abstinence and reduction data. Multiple imputation was performed using the fully conditional specification approach with 5 imputed data sets and results combined using the Rubin rules (eMethods in Supplement 2). Other prespecified sensitivity analyses examined the effect of imbalances in baseline participant characteristics using multiple logistic regression models to estimate odds ratios and 95% CIs [confidence intervals] for point prevalence abstinence at 12 and 24 weeks, adjusting for characteristics for which the absolute value of the standardized difference was 0.1 or greater. We conducted additional post hoc analyses: (1) to examine potential clustering by site using generalized linear mixed models with a random effect for site to estimate odds ratios and 95% CIs for point prevalence abstinence at 12 and 24 weeks, and (2) to compare the baseline characteristics of participants with self-reported smoking data at 12 weeks (primary end point) with those of participants without self-reported smoking data. Statistical analyses were performed using SAS statistical software (version 9.4; SAS Institute)."<sup>406</sup>

"Several prespecified sensitivity analyses were done. First, assessment of the effect of missing data on the primary outcome was done using multiple imputation by chained equations method (MICE). This imputation model included all the variables in the primary ITT [intention to treat] analysis, secondary outcomes (from each timepoint), and baseline variables associated with the missingness of the primary outcome. 20 imputed datasets were generated and combined using Rubin's rules, and the primary analysis model was then repeated using the imputed data. We specified a priori the following potential exploratory analyses to assess effect modification on the primary outcome: baseline hypertension, baseline MMSE [Mini-Mental State Examination], baseline age, time since Alzheimer's disease diagnosis, baseline brain volume, and change in systolic blood pressure. A post-hoc analysis was also done to investigate for differences between aggregated and disaggregated MRI [magnetic resonance imaging] data (according to MRI scanner modality) for the primary outcomes."<sup>407</sup>

"Four sensitivity analyses were done examining the primary outcome: restricted to women who had not received antibiotics in the 7 days before delivery, to examine whether any masking of a prophylactic effect was occurring by inclusion of pretreated women; excluding women prescribed antibiotics (other than the trial intervention) within the first 24 h after delivery, and who might therefore already have had an infection at the time of administration of the intervention; restricted to

women whose primary outcome was obtained between weeks 6 and 10 after delivery to exclude any biases by over-reporting of outcomes from data returned at a later timepoint or under-reporting of outcomes in data returned at an earlier timepoint; and including centre as a random effect. No subgroup analyses were planned; however, we did a post-hoc subgroup analysis of the primary outcome according to mode of birth (forceps or vacuum extraction). More stringent 99% CIs [confidence intervals] are presented for the estimate of RR [risk ratio] for this post-hoc subgroup analysis."<sup>408</sup>

"A prespecified subgroup analysis for the primary outcomes, testing for an interaction for baseline anxiety, depression, and opioid use, defined using their median values was completed. Prespecified sensitivity analyses for the primary outcome, excluding participants included in process evaluation interviews, adjusting for the imbalance of death, and split by baseline pain disorders were also completed. Because of the potential for type I error due to multiple comparisons, findings for analyses of secondary end points should be interpreted as exploratory. Statistical analyses were conducted using Stata version 16.1 (StataCorp)."<sup>409</sup>

*Explanation*

Sensitivity analyses can be important additional analyses to examine the robustness of the primary trial results under a range of assumptions about the data, methods, and models that differ from those of the primary analysis. When the findings from a sensitivity analysis are consistent with the primary trial findings, trialists can be confident that any assumptions in the primary analysis had little impact—strengthening the trial results. Morris and colleagues provide a principled approach to guide any sensitivity analyses by posing three questions to trialists: does the proposed sensitivity analysis address the same question as the primary analysis; is it possible for the proposed sensitivity analysis to return a different result to the primary analysis; and if the results do differ, is there any uncertainty as to which will be believed.<sup>402 410</sup>

Subgroup analyses are another set of additional analyses that are widely carried out and reported.<sup>411-414</sup> Here, the focus is on those analyses that look for evidence of a difference in treatment effect in complementary subgroups (eg, older and younger participants), a comparison known as a test of interaction.<sup>415</sup> Empirical analyses of subgroup difference claims for factors such as age, sex, race, ethnicity, and other factors show selective reporting, frequent lack of proper statistical support, and poor independent corroboration.<sup>416-418</sup>

A common but misleading approach is to compare P values for separate analyses of the treatment effect in each group. Categorising continuous variables to create subgroups is often done for simplicity and because it is perceived as easier to understand and communicate. Major limitations of the approach include the splitting of a continuous variable into discrete subgroups by arbitrarily chosen cut-off points that lack clinical or biological plausibility, which loses

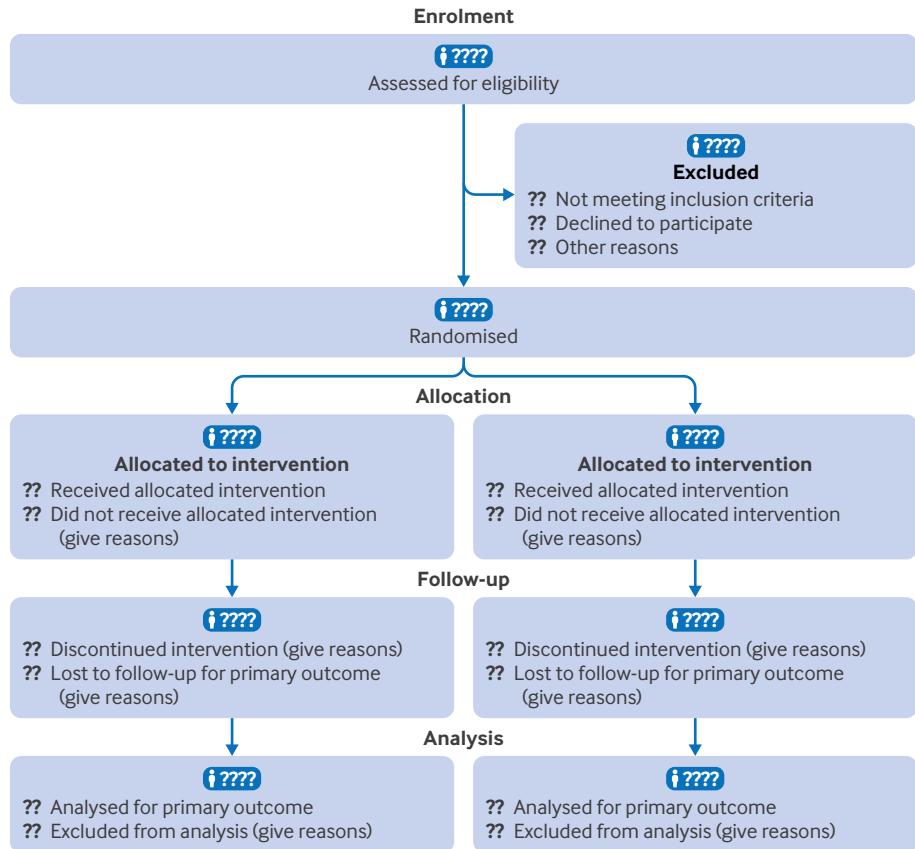


Fig 1 | CONSORT 2025 flow diagram. Flow diagram of participant progress through the phases of a two group, parallel randomised trial (ie, enrolment, intervention allocation, follow-up, and data analysis). CONSORT=Consolidated Standards of Reporting Trials

information, and thus reduces statistical power.<sup>419</sup> Choosing cut-off points based on achieving statistical significance should be avoided. It is incorrect to infer a subgroup effect (interaction) from one significant (in one subgroup) and one non-significant P value (in another subgroup). The rationale for any subgroups should be outlined (including how they are defined), along with whether the subgroups were specified a priori in the protocol or statistical analysis plan or were done post hoc. Because of the high risk for spurious findings, subgroup analyses are often discouraged. Post hoc subgroup comparisons (analyses done after looking at the data) are especially likely not to be confirmed by further studies. Most of these analyses do not have substantial credibility.

An alternative and stronger approach, which avoids the need to specify cut-off points to assess the interaction between a continuous variable (eg, age) and treatment effect would be to fit a regression model, which can be presented graphically to examine how the estimated treatment effects varies with the level of the variable.<sup>420</sup> These analyses are more complex, requiring model assumptions to capture the relationship (linear or non-linear) between the variable and the treatment effect. Authors should clearly describe the statistical methods used to explore the treatment-covariate interaction.

## CONSORT 2025: Results

**Item 22a:** For each group, the numbers of participants who were randomly assigned, received intended intervention, and were analysed for the primary outcome

### Examples

See figure 1, figure 2, and figure 3.

### Explanation

The design and conduct of some randomised trials are straightforward, and the flow of participants, particularly where there are no losses to follow-up or exclusions, through each phase of the study can be described relatively easily. For other trials, it can be difficult for readers to discern whether and why some participants did not receive the treatment as allocated, were lost to follow-up, or were excluded from the analysis.<sup>423</sup> This information is crucial for several reasons. Participants who were excluded after allocation are unlikely to be representative of all participants in the study. For example, participants may not be available for follow-up evaluation because they experienced an acute exacerbation of their illness or harms of treatment.<sup>271 424</sup>

Attrition as a result of loss to follow-up, which is often unavoidable, needs to be distinguished from investigator-determined exclusion for such reasons

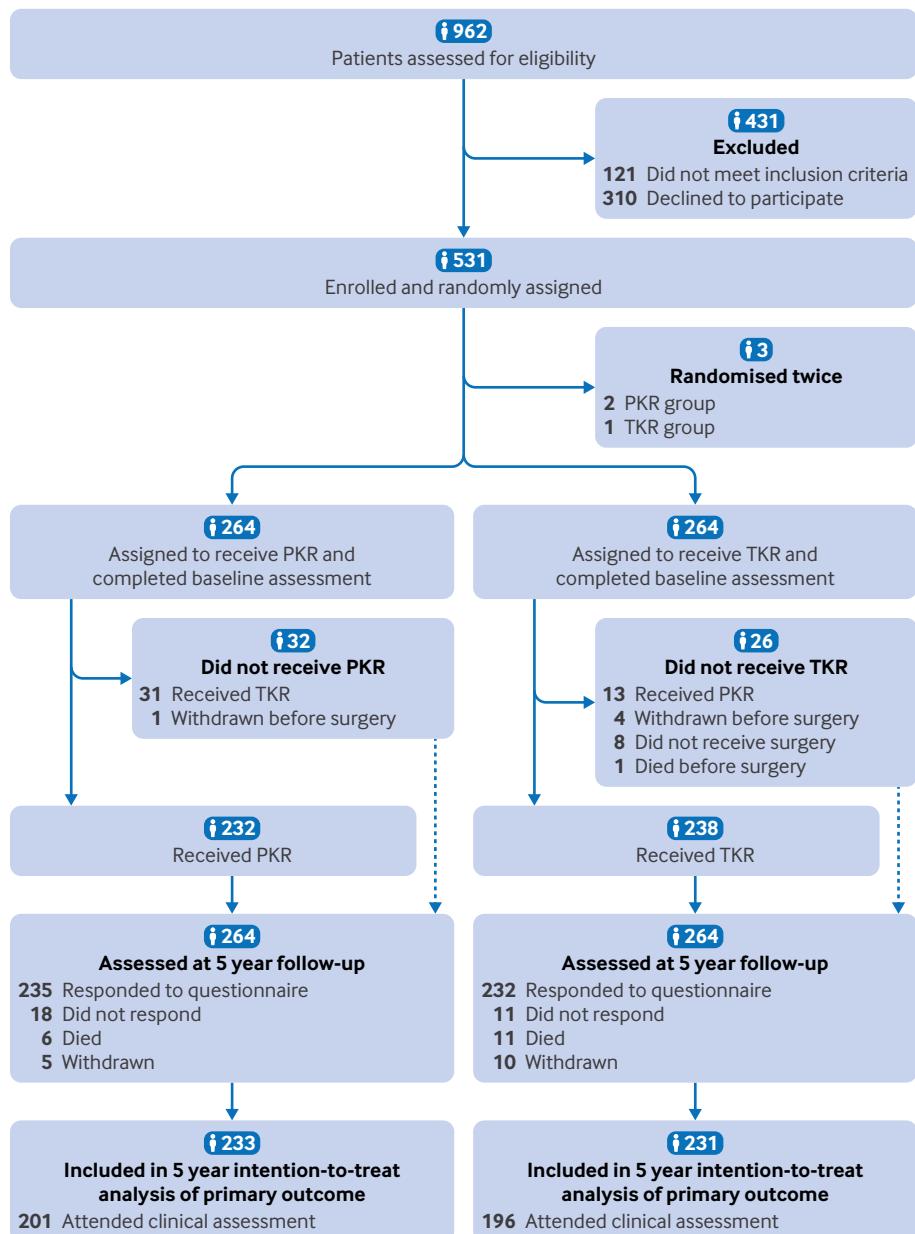
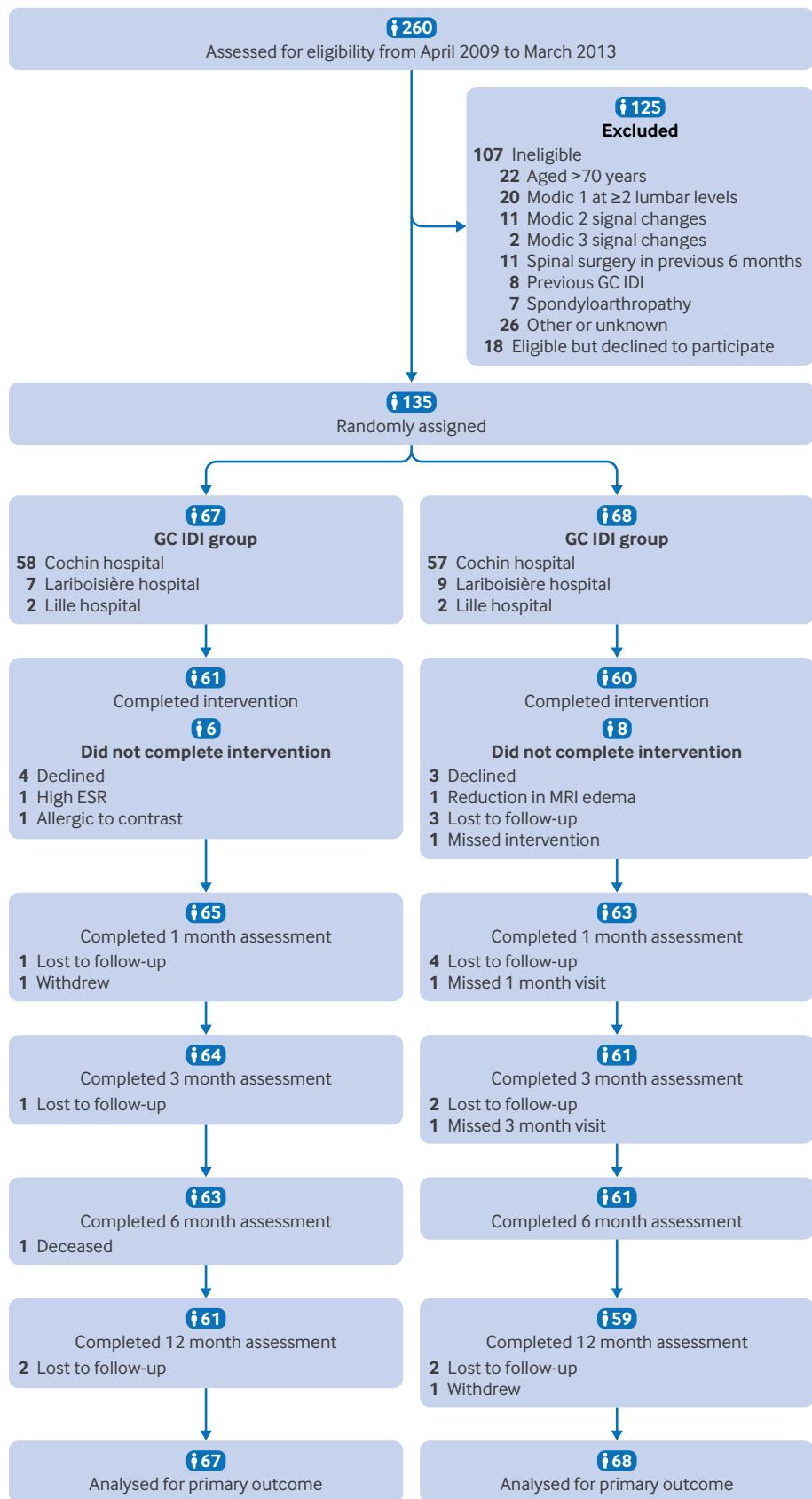


Fig 2 | Flow diagram of a multicentre trial of total (TKR) versus partial (PKR) knee replacement<sup>421</sup>

as ineligibility, withdrawal from treatment, and poor adherence to the trial protocol. Erroneous conclusions can be reached if participants are excluded from analysis, and imbalances in such omissions between groups may be especially indicative of bias.<sup>424-426</sup> Information about whether the investigators included in the analysis all participants who underwent randomisation, in the groups to which they were originally allocated (item 21b), is therefore of particular importance. Knowing the number of participants who did not receive the intervention as allocated or did not complete treatment permits the reader to assess to what extent the estimated efficacy of therapy might be underestimated in comparison with ideal circumstances.

If available, the number of people assessed for eligibility, and reason for exclusion, should also be reported. Although this number is relevant to external validity only and is arguably less important than the other counts,<sup>427</sup> it is a useful indicator of whether trial participants were likely to be representative of all eligible participants.

A suggested template for reporting the number of participants who were randomly assigned, received intended treatment, were lost to follow-up, and were analysed for the primary outcome is shown in figure 1, and the counts required are described in detail in table 7. A review of randomised trials published in general medical journals found that reporting of what happened to participants and their data was



**Fig 3 | Flow diagram of a multicentre trial of glucocorticoid intradiscal injection in patients with chronic low back pain.<sup>422</sup> ESR=erythrocyte sedimentation rate; GC IDI=glucocorticoid intradiscal injection; MRI=magnetic resonance imaging**

**Table 7 | Information required to document the flow of participants through each stage of a randomised trial for the primary outcome**

Stage	No of people included	No of people not included or excluded	Rationale
Enrolment	People evaluated for potential enrolment	People who did not meet the inclusion criteria or met the inclusion criteria but declined to be enrolled	These counts indicate whether trial participants were likely to be representative of all patients seen; they are relevant to assessment of external validity only, and they are often not available
Randomisation	Participants randomly assigned	—	Crucial count for defining trial size and assessing whether a trial has been analysed by intention-to-treat
Treatment allocation	Participants who received intervention as allocated, by trial group	Participants who did not receive intervention as allocated, by trial group	Important counts for assessment of internal validity and interpretation of results; reasons for not receiving intervention as allocated should be given
Follow-up	Participants who completed intervention as allocated, by trial group Participants who completed follow-up as planned, by trial group	Participants who did not complete intervention as allocated, by trial group Participants who did not complete follow-up as planned, by trial group	Important counts for assessment of internal validity and interpretation of results; reasons for not completing intervention or follow-up should be given
Analysis	Participants included in main analysis for the primary outcome, by trial group	Participants excluded from main analysis for the primary outcome, by trial group	Crucial count for assessing whether a trial has been analysed by intention-to-treat; reasons for excluding participants should be given

considerably more thorough in articles that included a diagram of the flow of participants through a trial than in those that did not.<sup>423</sup>

Some information, such as the number of individuals assessed for eligibility, may not always be known,<sup>428</sup> and depending on the nature of a trial, some counts may be more relevant than others. It will sometimes be useful or necessary to adapt the structure of the flow diagram to a particular trial. In some situations, other information may usefully be added. For example, for trials of non-pharmacological interventions it may be important to report the number of care providers or centres performing the intervention in each group and the number of participants treated by each care provider or in each centre.<sup>22</sup>

The exact form and content of the flow diagram may be varied according to specific features of a trial. For example, many trials of surgery or vaccination do not include the possibility of discontinuation. Although CONSORT strongly recommends using a flow diagram to communicate participant flow throughout the study, there is no specific, prescribed format.

#### Item 22b: For each group, losses and exclusions after randomisation, together with reasons

##### Example

See figure 2.<sup>421</sup>

##### Explanation

Some protocol deviations may be reported in the flow diagram (item 22a), for example, participants who did not receive the intended intervention. If participants were excluded after randomisation (contrary to the intention-to-treat principle) because they were found not to meet eligibility criteria, they should be included in the flow diagram. Use of the term “protocol deviation” in published articles is not sufficient to justify exclusion of participants after randomisation. The nature of the protocol deviation and the exact reason for excluding participants after randomisation should always be reported. Similarly, if participants did not complete the treatment intervention as allocated (ie, discontinued treatment), this should be reported by trial group, with reasons.

While attrition as a result of loss to follow-up is often unavoidable, it is important to report the number of participants who do not complete follow-up as planned; this should be reported by trial group along with reasons. For the primary outcome, loss to follow-up can be reported in the flow diagram.

#### Item 23a: Dates defining the periods of recruitment and follow-up for outcomes of benefits and harms

##### Examples

“Age-eligible participants were recruited . . . from February 1993 to September 1994 . . . Participants attended clinic visits at the time of randomization (baseline) and at 6-month intervals for 3 years.”<sup>429</sup>

“The trial involved five visits: Visit 1 on day 1 (screening, randomization and initial dosing), Visit 2 on day 2 (assessment of the primary endpoint), Visit 3 on day 4 (assessment of efficacy and safety parameters), Visit 4 on day 6 (end-of-treatment visit) and Visit 5 on day 8 to day 10 (follow up by telephone interview). Patients were asked to return all unused trial medication and their diaries at each visit.”<sup>430</sup> In this example, the term “safety” is used as a reference to harm outcomes; we recommend against the use of “safety”; preferable terms are “harms” or “adverse events.”

##### Explanation

Knowing when a study took place and over what period participants were recruited places the study in historical context. Medical and surgical treatments, including concurrent treatments, evolve continuously and may affect the routine care given to participants during a trial. Thus, it is important to report the periods of recruitment into the trial. Knowing the rate at which participants were recruited may also be useful, especially to other investigators.

The length of follow-up is not always a fixed period after randomisation. In many randomised trials in which the outcome is time to an event, follow-up of all participants is ended on a specific date. This date should be given, and it is also useful to report the minimum, maximum, and median duration of follow-up.<sup>431 432</sup> A review of reports in oncology journals

that used survival analysis, most of which were not randomised trials,<sup>432</sup> found that nearly 80% (104 of 132 reports) included the starting and ending dates for accrual of patients, but only 24% (32 of 132 reports) also reported the date on which follow-up ended.

Information on the periods of recruitment and follow-up may be different for outcomes of benefits and harm.<sup>20</sup> For example, the assessment of harms might be planned to take place during the entire study through non-systematic assessment, might occur during only part of the study duration, might occur at specific time points using systematic or non-systematic assessment, or might continue after the completion of follow-up for the main efficacy outcome. Reporting the periods of recruitment and follow-up for benefits and harms is important to allow comprehensive and accurate interpretation of the trial's results.<sup>20</sup>

#### **Item 23b: If relevant, why the trial ended or was stopped**

##### *Examples*

"At the time of the interim analysis, the total follow-up included an estimated 63% of the total number of patient-years that would have been collected at the end of the study, leading to a threshold value of 0.0095, as determined by the Lan-DeMets alpha-spending function method . . . At the interim analysis, the RR [risk ratio] was 0.37 in the intervention group, as compared with the control group, with a p value of 0.00073, below the threshold value. The Data and Safety Monitoring Board advised the investigators to interrupt the trial and offer circumcision to the control group, who were then asked to come to the investigation centre, where MC (medical circumcision) was advised and proposed . . . Because the study was interrupted,

some participants did not have a full follow-up on that date, and their visits that were not yet completed are described as 'planned' in this article."<sup>433</sup>

"In January 2000, problems with vaccine supply necessitated the temporary nationwide replacement of the whole cell component of the combined DPT/Hib vaccine with acellular pertussis vaccine. As this vaccine has a different local reactogenicity profile, we decided to stop the trial early."<sup>434</sup>

##### *Explanation*

Arguably, trialists who arbitrarily conduct unplanned interim analyses after very few events accrue using no statistical guidelines run a high risk of catching the data at a random extreme, which likely represents a large overestimate of treatment benefit.<sup>435</sup>

Readers will likely draw weaker inferences from a trial that was truncated in a data driven manner versus one that reports its findings after reaching a results-independent goal (box 9). Where relevant, authors should report the reason for stopping the trial before completion as planned (eg, result of an interim analysis, lack of funding, poor recruitment of participants, intervention no longer available, or the question becoming no longer relevant after publication of another study). Authors should also disclose factors extrinsic to the trial that affected the decision to stop the trial, and who made the decision to stop the trial, including reporting the role the funding agency had in the deliberations and in the decision to stop the trial.<sup>436</sup>

A systematic review of 143 randomised trials that were stopped earlier than planned for benefit found that these trials reported stopping after accruing a median of 66 events. The review estimated a median relative risk of 0.47 and a strong association between

#### **Box 9: Early stopping of randomised trials**

Randomised trials can end when they reach their sample size goal, their event count goal, their length of follow-up goal, or their scheduled date of closure. In these situations, the trial will stop in a manner independent of its results and stopping is unlikely to introduce bias in the results. Alternatively, randomised trials can stop earlier than planned because of the result of an interim analysis showing larger than expected benefit or harm of the experimental intervention. Randomised trials can also stop earlier than planned when investigators find evidence of no important difference between experimental and control interventions (ie, stopping for futility). In addition, trials may stop early because the trial becomes unviable: funding vanishes, researchers cannot access eligible patients or study interventions, or the results of other studies make the research question irrelevant.

Full reporting of why a trial ended is important for evidence based decision making (item 23b). Researchers<sup>436</sup> examining why 143 trials stopped early for benefit found that many failed to report key methodological information regarding how the decision to stop was reached: the planned sample size (n=28), interim analysis after which the trial was stopped (n=45), or whether a stopping rule informed the decision (n=48). Item 16b of the CONSORT checklist requires the reporting of timing of interim analyses, what triggered them, how many took place, whether these were planned or ad hoc, and whether there were statistical guidelines and stopping rules in place a priori. Furthermore, it is helpful to know whether an independent data monitoring committee participated in the analyses (and who composed it, with particular attention to the role of the funding source), and who made the decision to stop. Often the data monitoring committee make recommendations and the funders (sponsors) or the investigators make the decision to stop.

Trials that stop early for reasons apparently independent of trial findings, and trials that reach their planned termination, are unlikely to introduce bias by stopping.<sup>437</sup> In these cases, the authors should report whether interim analyses took place and whether these results were available to the funder.

The push for trials that change the intervention in response to interim results, thus enabling a faster evaluation of promising interventions for rapidly evolving and fatal conditions, will require even more careful reporting of the process and decision to stop trials early.<sup>174 438</sup>

CONSORT=Consolidated Standards of Reporting Trials.

the number of events accrued and the magnitude of the effect, with smaller trials with fewer events yielding the largest treatment effects (odds ratio 31, 95% CI 12 to 82).<sup>436</sup> While an increasing number of trials published in high impact medical journals report stopping early, many still do not report how the decision to stop the trial was made. In a systematic review of 110 paediatric trials that reported on the presence of a data monitoring committee, interim analysis, or early stopping, 32 were terminated early. Of these 32 trials, 22 (69%) did not report predefined stopping guidelines and 15 (47%) did not provide information on statistical monitoring methods.<sup>439</sup>

**Item 24a: Intervention and comparator as they were actually administered (eg, where appropriate, who delivered the intervention/comparator, whether participants adhered, whether they were delivered as intended (fidelity))**

*Examples*

“Patients were randomly assigned to the P2Y12 inhibitor monotherapy group (aspirin plus a P2Y12 inhibitor for 3 months and thereafter a P2Y12 inhibitor alone) or to the DAPT group (aspirin plus a P2Y12 inhibitor for at least 12 months) in a 1:1 ratio... Overall adherence to the study protocol was 79.3% in the P2Y12 inhibitor monotherapy group and 95.2% in the DAPT group... The rates of P2Y12 inhibitor use were similar in both groups: 96.4% at 6 months and 95.0% at 12 months in the P2Y12 inhibitor monotherapy group and 98.1% at 6 months and 96.6% at 12 months in the DAPT group. The median duration of aspirin was 96 days (interquartile range, 88–118 days) in the P2Y12 inhibitor monotherapy group and 365 days (interquartile range, 363–365) in the DAPT group. The proportion of patients receiving aspirin beyond 3 months in the P2Y12 inhibitor monotherapy group was 14.4% at 6 months and 8.9% at 12 months.”<sup>440</sup>

“Most participants received treatment as allocated [table 8]. Across intervention groups, high protocol adherence was achieved in terms of the delivery, type, and content for the injection, progressive exercise, and best practice advice interventions. 53 physiotherapists delivered corticosteroid injections to 329 (97%) participants and three doctors to ten

(3%) participants. Progressive exercise was delivered by 104 physiotherapists to 339 participants and best practice advice was delivered by 83 physiotherapists to 324 participants. Two physiotherapists swapped groups during the trial because of staffing issues and delivered both interventions. We found no difference in attendance rates between those receiving progressive exercise or best practice advice and those who received the intervention in conjunction with corticosteroid injection [table 8].”<sup>114</sup>

*Explanation*

This new item has been added to the CONSORT 2025 checklist to address the poor reporting of the intervention and comparator in randomised trials.<sup>195 198–201 214</sup> For example, in a review of 102 randomised trials evaluating bariatric surgery, only 14% reported the intervention as implemented.<sup>441</sup> A review of 192 randomised trials assessing pharmacological treatments in six major chronic diseases published in journals with high impact factors showed that adherence to medication was reported in only one third of the publications.<sup>442</sup> A review of 100 randomised trial reports published in general medical journals with high impact factors highlighted that only 11% of the trials assessing long term interventions and 38% of those assessing short term interventions adequately reported treatment initiation and completeness of treatment.<sup>209</sup> A review of 111 randomised trials reports showed that only 46% reported adherence results.<sup>443</sup> A review of 94 placebo/sham controlled randomised trials published in high impact journals showed that only 54% reported actual adherence or fidelity.<sup>214</sup>

There is frequently a gap between the intervention/comparator as planned and described in the trial protocol and how the intervention/comparator were actually administered. This gap could be related to poor fidelity, which can be driven by the extent to which the intervention/comparator are implemented as planned in the protocol by practitioners, and/or poor adherence to treatment, defined as the extent to which participants comply with the care providers’ recommendations (eg, taking a drug, placebo, behavioural change, doing exercises).<sup>209 444</sup> This gap

**Table 8 | Example of good reporting: Intervention received by treatment group**

	Best practice advice (n=174)	Injection and best practice advice (n=178)	Progressive exercise (n=174)	Injection and progressive exercise (n=182)
Injection received	—	168 (94)	—	171 (94)
Injection not received with reasons	—	10 (6)	—	11 (6)
Received extra injection	—	0	—	2 (1)
Completed exercise treatment	162 (93)	162 (91)	138 (79)	139 (76)
Partial exercise completion	—	—	29 (17)	33 (18)
Median (IQR) number of sessions	1 (1–1)	1 (1–1)	4 (3–6)	4 (3–5)
Completed session 1	162 (93)	162 (91)	167 (96)	172 (95)
Completed session 2	—	—	161 (93)	160 (88)
Participants who received additional sessions	3 (2)	5 (3)	3 (2)	2 (1)

Data are number (%) or participants unless stated otherwise. Adapted from Hopewell et al, 2021<sup>114</sup>  
IQR=interquartile range.

could also be related to the expected diversity in the implementation of the intervention/comparator in clinical practice particularly for complex interventions.

The gap between the intervention/comparator as planned and as delivered also depends on how the trial was planned. In explanatory trials, the aim is to estimate treatment effect under ideal circumstances. The intervention/comparator are usually highly standardised with close monitoring of fidelity and adherence to interventions and strategies to increase them. Intensive efforts to maximise fidelity and adherence in early phase trials or explanatory trials may lead to unrealistic, inflated estimates of treatment benefit that cannot be reproduced under real life circumstances.<sup>445 446</sup> Reporting the results of this monitoring is essential to allow readers to interpret the study results.

In contrast, pragmatic trials aim to determine treatment effect in clinical conditions. The intervention and comparator are usually highly flexible, and measurement of fidelity and adherence is unobstructive with no strategies to maintain or improve them.

Reporting how the intervention/comparator were actually administered is nevertheless crucial to allow readers to accurately interpret the trial results. For example, in a large international randomised trial comparing endarterectomy to medical management for patients with symptomatic carotid artery stenosis, there were important differences in the delay in receiving the surgical procedure which impacted the outcomes.<sup>179</sup>

Authors should provide details on who actually delivered the intervention/comparator (number and expertise), how the intervention/comparator were delivered, what was actually administered, participants' adherence to treatment, and the caregiver's fidelity to the intervention/comparator protocol where appropriate. Reporting fidelity and adherence can be complex and vary according to the type of intervention or comparator (eg, one-off, short term repeated, long term repeated). Various deviations to the protocol can occur. Participants might initiate the intervention/comparator but then discontinue the intervention/comparator permanently and completely after a specific period of time, discontinue temporarily, reduce the dose, or modify the schedule.

More detailed information is available in TIDie<sup>23</sup> and the CONSORT extension for non-pharmacological treatments<sup>22</sup> (item 13).

#### **Item 24b: Concomitant care received during the trial for each group**

##### *Example*

"The principal investigators invited hospitals with the capability to provide the current standard of care for covid-19 to participate in the study. Minimum requirements for the standard of care included the provision of intravenous fluids, supplemental oxygen, regular laboratory testing, SARS-CoV-2 testing, haemodynamic monitoring, and intensive care, as well as the ability to deliver concomitant medications  
..."

"90 (60%) patients received concomitant drug treatment before randomisation. Among these, 52 (35%) patients received antiviral treatment [table 9]. . . . Concomitant treatments, including antiviral agents, antibiotics, and systemic glucocorticoid therapy, were similar in the two groups [table 9]."<sup>447</sup>

##### *Explanation*

Concomitant care refers to any additional treatments, interventions, or medications that participants may have received during the trial period, in addition to the study interventions. Relevant concomitant care refers to interventions that could have affected the outcome. transparently reporting this information is essential for readers to understand the context in which the trial was conducted and be able to assess the potential influence of concomitant care on the trial's results. Particularly, readers should be aware of any unequal use of concomitant care that might affect the outcome between the intervention and comparator groups.<sup>448</sup> This is particularly important when the trials are not fully blinded.<sup>449</sup> This information could be particularly important for evaluating the risk of bias due to deviations from the intended interventions, an important domain of the risk-of-bias tool developed by Cochrane.<sup>210 211</sup>

Nevertheless, this information is poorly reported. A review of 164 cardiovascular clinical trials published in five influential medical journals from 2011 to 2021 showed that cointerventions were inadequately reported in 71% of the trials<sup>450</sup> and that trials with deficient reporting had larger treatment effect estimates on average.<sup>451</sup> In rheumatology, an assessment of 109 trials in leading journals from 2018 to 2020 found that only 57% of randomised trials provided the number of patients on concomitant medications at baseline, and only 5% reported the cumulative or mean exposure data for concomitant medications.<sup>452</sup>

Authors should report the number and percentage of participants receiving the different relevant concomitant care in each arm and, where relevant, the cumulative or average for each concomitant intervention taken over the trial period for each group.

#### **Item 25: A table showing baseline demographic and clinical characteristics for each group**

##### *Example*

See table 10.<sup>453</sup>

##### *Explanation*

Although the eligibility criteria (item 12a) indicate who was eligible for the trial, it is also important to know the characteristics of the participants who were actually included. This information allows readers, especially clinicians, to judge how relevant the results of a trial might be to an individual patient. Participant baseline demographics may include characteristics such as age, sex and/or gender,<sup>182</sup> place of residence, race and/or ethnicity, culture and/or religion, language, occupation, education, or socioeconomic status. Baseline clinical characteristics include those

**Table 9 | Example of good reporting: Treatments after randomisation in patients in an intention-to-treat population**

Drug treatment after randomisation	Standard of care plus hydroxychloroquine (n=75)	Standard of care (n=75)	Total (n=150)
Antiviral agents	47 (63)	48 (64)	95 (63)
Arbidol	37 (49)	33 (44)	70 (47)
Virazole	13 (17)	15 (20)	28 (19)
Lopinavir-ritonavir	13 (17)	12 (16)	25 (17)
Oseltamivir	8 (11)	9 (12)	17 (11)
Entecavir	1 (1)	1 (1)	2 (1)
Antibiotics	32 (43)	27 (36)	59 (39)
Systemic glucocorticoid treatment	6 (8)	4 (5)	10 (7)

Data are number (%). Adapted from Tang et al.<sup>447</sup>

which are identical, or closely related, to the trial outcomes.<sup>454</sup>

Randomised trials aim to compare groups of participants that differ only with respect to the intervention (treatment). Although proper random assignment prevents selection bias, it does not guarantee similarity of the groups at baseline. Any differences in baseline characteristics are, however, the result of chance rather than bias.<sup>294</sup> Important demographic and clinical characteristics should be presented so that readers can assess how similar the groups were at baseline. Baseline data are especially valuable for outcomes that can also be measured at the start of the trial (eg, blood pressure).

Baseline information is most efficiently presented in a table. For continuous variables, such as weight

or blood pressure, the variability of the data should be reported, along with average values. Continuous variables can be summarised for each group by the mean and standard deviation. When continuous data have an asymmetrical distribution, a preferable approach may be to quote the median and percentile values (eg, the 25th and 75th percentiles).<sup>365</sup> Standard errors and CIs are not appropriate for describing variability—they are inferential rather than descriptive statistics. Variables with a small number of ordered categories (such as stages of disease I to IV) should not be treated as continuous variables; instead, numbers and proportions should be reported for each category.<sup>364 365</sup>

Significance testing of baseline differences is not recommended and should not be reported.<sup>269 294 455</sup>

**Table 10 | Example of good reporting: Baseline demographic and clinical characteristics between study groups**

Characteristic	Usual care (n=242)	Exercise (n=246)
Mean (SD) age (years)	63.5 (11)	61.3 (12)
Female sex	186 (76)	188 (76)
Ethnic origin		
White	235 (98)	238 (97)
Indian	2 (1)	3 (1)
Pakistani	1 (<1)	—
Mixed	1 (<1)	3 (1)
Other	1 (<1)	2 (1)
Dominant in right hand	215 (90)	226 (92)
Median (IQR) No of years since rheumatoid arthritis diagnosis, estimated by participant	10 (4-22)	10 (4-21)
Median (IQR) baseline ESR	16 (8-28)	15 (7-28)
Median (IQR) baseline CRP	6 (3-12)	5 (3-12)
Drug treatment		
Biological DMARD	52 (22)	51 (21)
Combination non-biological DMARD	53 (22)	72 (29)
Single non-biological DMARD	118 (49)	103 (42)
Other drugs	19 (8)	19 (8)
Mean (SD) MHQ		
Overall hand function (both)	52.1 (16.4)	52.1 (15.2)
Activities of daily living (both)	54.1 (25.0)	54.5 (24.5)
Work	48.4 (22.0)	48.2 (22.0)
Pain	51.4 (19.9)	51.9 (21.9)
Aesthetics (both)	58.6 (22.1)	56.9 (22.0)
Satisfaction (both)	43.5 (22.3)	43.9 (19.7)
Overall score	50.9 (16.9)	50.6 (16.4)
Mean (SD) SF-12 score		
Aggregate physical scale (PCS)	34.5 (9.5)	33.8 (9.8)
Aggregate mental scale (MCS)	48.9 (11.0)	48.1 (10.7)

Data are number (%) of participants unless stated otherwise. DMARD=disease-modifying drugs; ESR=erythrocyte sedimentation rate; IQR=interquartile range; MHQ=Michigan Hand Outcome Questionnaire; PCS=physical component score; MCS=mental component score; SD=standard deviation. Adapted from Lamb et al.<sup>453</sup>

Such significance tests assess the probability that observed baseline differences could have occurred by chance; however, providing the randomisation has not been subverted or comprised, any differences are caused by chance. Unfortunately, such significance tests are still relatively common.<sup>456-458</sup> Such hypothesis testing is superfluous and can mislead investigators and their readers.<sup>459</sup> Rather, comparisons at baseline should be based on consideration of the prognostic strength of the variables measured and the magnitude of any chance imbalances that have occurred.<sup>459</sup>

#### Item 26: Items for each primary and secondary outcome

For each primary and secondary outcome, by group:

- the number of participants included in the analysis
- the number of participants with available data at the outcome time point
- result for each group, and the estimated effect size and its precision (such as 95% CI)
- for binary outcomes, presentation of both absolute and relative effect size.

#### Examples

“All principal analyses were based on the intention-to treat (ITT) principle, analysing participants in the groups to which they were randomly assigned irrespective of compliance with treatment allocation.”<sup>460</sup>

See table 11,<sup>123</sup> table 12,<sup>123</sup> and table 13.<sup>461</sup>

#### Explanation

For each primary and secondary outcome, the number of participants included in each group is an essential element of the analyses. Although the flow diagram (item 22a) should indicate the numbers of participants included in the analysis of the primary outcome, the number of participants with available data will often vary for different outcomes and at different time points.

Missing data can introduce potential bias through different types of participants being included in each treatment group. It can also reduce, through

loss of information, the power to detect a difference between treatment groups if one exists (item 21c) and reduce the generalisability of the trial findings.<sup>404</sup> It is therefore important to report the number of participants with available data for each primary and secondary outcome and at each timepoint. Where possible, it is also important to report the reason data were not available, for example, if the participant did not attend follow-up appointments, or if data were truncated because the participant died.<sup>404</sup> The extent and causes of missing data can vary. For example, a systematic review of palliative care trials estimated that 23% of primary outcome data were not available;<sup>462</sup> this compares to a recent review of trials published in four top general medical journals where the median percentage of participants with a missing outcome was around 9%.<sup>392</sup>

Trial results are often more clearly displayed in a table rather than in the text, as shown in table 11 and table 12. For each outcome, results should be reported as a summary of the outcome in each group (eg, the number of participants included in the analysis with or without the event and the denominators, or the mean and standard deviation of measurements), together with the contrast between the groups, known as the effect size. For binary outcomes, the effect size could be the risk ratio (relative risk), odds ratio, or risk difference; for survival time data, it could be the hazard ratio or difference in median survival time; and for continuous data, it is usually the difference in means.

For all outcomes, authors should provide a CI to indicate the precision (uncertainty) of the estimated effect size.<sup>364, 463</sup> A 95% CI is conventional, but occasionally other levels are used. Most journals require or strongly encourage the use of CIs.<sup>464</sup> They are especially valuable in relation to differences that do not meet conventional statistical significance, for which they often indicate that the result does not rule out an important clinical difference. The use of CIs has increased markedly in recent years, although not in all medical specialties.<sup>465</sup> A common error is the presentation of separate CIs for the outcome in

**Table 11 | Example of good reporting: Summary results for each trial group (binary outcomes). Secondary outcomes of arm symptoms, and lymphoedema by treatment group. Data are number (%) unless stated otherwise**

Outcome	Usual care	Exercise	Adjusted odds ratio (95% CI)
<b>Acute and chronic postoperative pain*</b>			
Moderate to severe, 6 weeks	46/150 (31)	28/153 (18)	1.90 (1.02 to 3.52)
Moderate to severe, 6 months	30/133 (23)	25/145 (17)	1.42 (0.72 to 2.84)
Moderate to severe, 12 months	43/139 (31)	22/135 (16)	2.41 (1.24 to 4.70)
<b>Neuropathic pain, DN4 positive</b>			
6 weeks	21/150 (14)	24/153 (16)	0.73 (0.22 to 2.45)
6 months	29/133 (22)	26/145 (18)	1.64 (0.63 to 4.23)
12 months	32/139 (23)	22/135 (16)	1.29 (0.45 to 3.69)
<b>Lymphoedema, LBCQ</b>			
6 weeks	20/150 (13)	22/153 (14)	1.07 (0.52 to 2.24)
6 months	32/133 (24)	29/145 (20)	0.82 (0.43 to 1.56)
12 months	36/139 (26)	33/135 (24)	1.17 (0.62 to 2.23)

DN4=Douleur Neuropathique-4 (positive neuropathic pain=score >3); LBCQ=Lymphoedema and Breast Cancer Questionnaire (positive symptoms=arm both heavy and swollen). Adapted from Bruce et al.<sup>123</sup>

\*Numerical rating scale: acute and chronic post-operative pain. Moderate to severe pain=4-10.

**Table 12 | Example of good reporting: Summary results for each trial group (continuous outcomes). Disability of Arm, Shoulder, Hand (DASH) scores by treatment group**

Time point, analysis	Usual care		Exercise		Difference between groups (95% CI)*	
	No	Mean (SD)	No	Mean (SD)	Unadjusted	Adjusted
6 months, ITT	125	20.8 (20.1)	134	18.0 (17.1)	2.76 (-1.79 to 7.31)	4.60 (0.30 to 8.90)
12 months, ITT (primary outcome)	138	23.7 (22.9)	132	16.3 (17.6)	7.34 (2.44 to 12.23)	7.81 (3.17 to 12.44)

Scores adjusted for age, baseline DASH, breast surgery, axillary surgery, radiotherapy, and chemotherapy. Mean differences in upper limb function favour exercise intervention. Adapted from Bruce et al.<sup>123</sup>

CI=confidence interval; ITT=intention to treat; SD=standard deviation.

\*Absolute mean difference between treatment groups.

each group rather than for the treatment effect.<sup>465</sup> Although P values may be provided in addition to CIs, results should not be reported solely as P values.<sup>466 467</sup> Results should be reported for all planned primary and secondary outcomes and at each time point, not just for analyses that were statistically significant or thought to be interesting. Selective reporting within studies is a widespread and serious problem.<sup>51 468</sup>

When the primary outcome is binary, both the relative effect (risk ratio (relative risk) or odds ratio), and the absolute effect (risk difference) should be reported (with CIs) (table 13), as neither the relative measure nor the absolute measure alone gives a complete picture of the effect and its implications. Different audiences may prefer either relative or absolute risk, but both clinicians and lay people tend to overestimate the effect when it is presented solely in terms of relative risk.<sup>469-471</sup> The magnitude of the risk difference is less generalisable to other populations than the relative risk since it depends on the baseline risk in the unexposed group, which tends to vary across populations. For diseases where the outcome is common, a relative risk near unity might nonetheless indicate clinically important differences in public health terms. In contrast, a large relative risk when the outcome is rare may not be so important for public health (although it may be important to an individual in a high risk category). For both binary and survival time data, expressing the results also as the number needed to treat for benefit or harm can be helpful.<sup>472 473</sup>

#### Item 27: All harms or unintended events in each group

##### Example

"Few women vomited after drug administration. 12 (0.2%) of 6685 sulfadoxine–pyrimethamine, 19 (0.3%) of 7014 dihydroartemisinin–piperaquine, and 23 (0.3%) of 6849 dihydroartemisinin–piperaquine plus azithromycin treatment courses were vomited within 30 min [table 14]. One (0.1%) of 1552

women in the sulfadoxine–pyrimethamine group, two (0.1%) of 1558 in the dihydroartemisinin–piperaquine group, and four (0.3%) of 1556 in the dihydroartemisinin–piperaquine plus azithromycin group vomited after their first course of treatment (the only course when azithromycin was coadministered with dihydroartemisinin–piperaquine; table 14). All three regimens were well tolerated ([table 14] . . .), but vomiting, nausea, and dizziness were more common in the first 3 days after dihydroartemisinin–piperaquine (13 [3.2%), 14 [3.4%), and 15 [3.7%) of 410 women visited at home, respectively) than sulfadoxine–pyrimethamine (one [0.3%), one [0.3%), and zero [0%) of 384 women visited at home, respectively; appendix pp 20–21). The addition of azithromycin to dihydroartemisinin–piperaquine was associated with significantly more vomiting than with dihydroartemisinin–piperaquine alone ( $p=0.0033$ ; [table 14])."<sup>196</sup>

##### Explanation

Readers need information about the harms as well as the benefits of interventions to make rational and balanced decisions. Randomised trials offer an excellent opportunity for providing harms data, although they cannot detect differences in uncommon or rare harms between treatment groups. The existence and nature of adverse effects can have a major impact on whether a particular intervention will be deemed acceptable and useful. Not all reported adverse events observed during a trial are necessarily a consequence of the intervention; some may be a consequence of the condition being treated. Nevertheless, they all need to be reported.

Many reports of randomised trials provide inadequate information on harms. A comparison between harm data submitted to the trials database of the National Cancer Institute, which sponsored the trials, and the information reported in journal articles found that low grade adverse events were

**Table 13 | Example of good reporting: Absolute and relative effect sizes**

Primary outcome	Early administration (%) (n=1344)	Delayed selective administration (%) (n=1346)	Risk ratio (95% CI)	Risk difference (95% CI)
Death or oxygen dependence at "expected date of delivery"	31.9 (429)	38.2 (514)	0.84 (0.75 to 0.93)	-6.3 (-9.9 to -2.7)

Data are % (number) of participants, unless stated otherwise. Adapted from Sola et al.<sup>461</sup> "The risk of oxygen dependence or death was reduced by 16% (95% CI 25% to 7%). The absolute difference was -6.3% (95% CI -9.9% to -2.7%); early administration to an estimated 16 babies would therefore prevent 1 baby dying or being long-term dependent on oxygen."<sup>461</sup>

CI=confidence interval.

**Table 14 | Example of good reporting: Safety and tolerability endpoints (incidence measures, events (incidence per person year at risk) or prevalence. Data are number (%) unless stated otherwise**

Endpoint	Treatment		
	Sulfadoxine-pyrimethamine	Dihydroartemisinin-piperaquine	Dihydroartemisinin-piperaquine plus azithromycin
<b>Adverse events</b>			
Dizziness	4 (0.8)	51 (9.5)	38 (7.2)
Vomiting	4 (0.8)	41 (7.7)	71 (13.5)
Nausea	2 (0.4)	44 (8.2)	35 (6.7)
Abdominal pain	6 (1.1)	11 (2.1)	14 (2.7)
Diarrhoea	2 (0.4)	5 (0.9)	6 (1.1)
Headache	12 (2.3)	17 (3.2)	18 (3.4)
Rash	0 (0.0)	2 (0.4)	0 (0.0)
<b>Serious adverse events and grade 3-4 adverse events (in pregnant women)</b>			
Any	95 (17.7)	79 (14.8)	92 (16.9)
Maternal mortality*	1/1553 (0.1)	2/1561 (0.1)	3/1557 (0.2)
<b>By system organ class</b>			
Blood and lymphatic system disorders	2 (0.4)	0 (0.0)	2 (0.4)

Adapted from Madanitsa et al.<sup>196</sup>

\*Number (%)/total number.

under-reported in journal articles. High grade events (Common Toxicity Criteria grades 3 to 5) were reported inconsistently in the articles and the information regarding attribution to investigational drugs was incomplete.<sup>474</sup> Moreover, a review of trials published in six general medical journals in 2006 to 2007 found that while 89% of 133 reports mentioned adverse events, no information on severe adverse events and withdrawal of patients owing to an adverse event was given in 27% and 48% of articles, respectively.<sup>475</sup> In a later review of 196 randomised trials of invasive pain treatments published in six major journals, 76% provided the denominators for analyses on harms and 85% reported the absolute risk per arm and per adverse event type, grade, and seriousness, and presented appropriate metrics.<sup>476</sup>

For non-systematically assessed harms, reporting can be more complex as the information is not standardised. A common approach is to code the event declared by participants. Authors should report the coding system used, whether coding was prespecified in the protocol, in the statistical analysis plan, or post hoc, and whether coding was performed by researchers blinded to the treatment allocated. In addition, there is a risk of under-reporting and selective non-reporting of harms particularly for non-systematically assessed harms. A reanalysis of individual participant data from

six randomised trials of gabapentin found evidence of important harms that were not disclosed in the published reports but identified after data sharing and reanalysis.<sup>477</sup> Sharing of de-identified individual participant data may be needed to be able to adequately synthesise this information, for example for inclusion in a systematic review.<sup>231 235 236</sup>

Authors should report for each group, the number of participants at risk, the number of deaths, the number of participants withdrawn due to harms, the number of participants with at least one harm event, and the number of events, if appropriate. Where appropriate, the estimated effect size with its precision (such as 95% CIs) should be reported including both absolute and relative effects for binary outcomes. It is important to separate the reporting of systematically and non-systematically assessed harms. Systematically assessed harms should be reported even if zero events were identified. It should also be clear whether the authors are reporting the number of participants with at least one harm event or the number of events per unit of time at risk and whether recurrent events were included. The number of participants withdrawn because of harms should also be reported for each group. Finally, results should be reported for all harms. We strongly discourage the use of thresholds or criteria to select which harms should be reported. All harms could be detailed in supplementary materials.

We recommend reporting the results in a table with the results for each trial arm.<sup>478</sup> More detailed information can be found in the CONSORT statement extension for harms, which was updated in 2022.<sup>20</sup>

#### Item 28: Any other analyses performed, including subgroup and sensitivity analyses, distinguishing prespecified from post hoc Examples

"In a [prespecified] sensitivity analysis to support the primary binary endpoint, the NRS [numerical

**Table 15 | Example of good reporting: Main, secondary, and subgroup analyses of Oxford Finger Nail Appearance Score cosmetic outcome<sup>479</sup>**

	Nail replaced	Nail discarded	Effect size*
<b>Main analysis</b>			
OFNAs, median (IQR)	5 (4-5)	5 (4-5)	0.55 (0.49 to 0.60)t
<b>Subgroup analyses</b>			
Assessor (parent v child)			0.24, (0.06 to 0.96)‡
Preoperative antibiotic use			1.11 (0.62 to 2.31)‡

Adapted from Jain et al.<sup>479</sup>

IQR=interquartile range; OFNAs=Oxford Finger Nail Appearance Score.

\*Values in parentheses are 95% confidence intervals; effect sizes are shown as odds ratios, except tprobability that OFNAs in discard arm is greater than that in replace arm from Mann-Whitney U test.

†Adjusted model allowed for intrasite correlation using cluster-robust standard errors.

rating score] pain score at 1 month was also analyzed using the constrained longitudinal data analysis model . . . Primary Outcome: At 1 month after the intervention, the percentage of responders (Low Back Pain intensity <40) was higher in the glucocorticoid intradiscal injection (GCIDI) group (36 of 65 [55.4%]) than the control group (21 of 63 [33.3%]) (absolute risk difference, 22.1 percentage points [CI, 5.5 to 38.7 percentage points];  $P=0.009$  [after multiple imputation]) . . . In the sensitivity analysis, the mean reduction in LBP [low back pain] intensity from baseline to 1 month was greater in the GCIDI group (-32.5 [CI, -38.2 to -26.8]) than the control group (-17.5 [CI, -23.3 to -11.7]) (absolute difference, -15.0 [CI, -22.9 to -7.1];  $P<0.001$ ).<sup>422</sup>

"Owing to the later inclusion of parent cosmetic appearance assessments (to assist with trial conduct), it was decided to perform a post hoc subgroup analysis to determine whether the scores given by the assessors and parents differed between treatment groups [table 15] . . . The assessor scores did not indicate a difference between the nail-replaced and nail-discarded groups. However, the scores given by the parents suggested that there was a statistically significant difference in favour of the nail-discarded group. The treatment by subgroup interaction term was statistically significant (OR [odds ratio] 0.24, 95% CI [confidence interval] 0.06 to 0.96.  $P=0.044$ )."<sup>422</sup>

#### *Explanation*

Multiple analyses of the same data create a risk for false-positive findings.<sup>480</sup> Authors should especially resist the temptation to perform many subgroup analyses.<sup>481-483</sup> Analyses that were prespecified in the trial protocol (item 3) are much more reliable than those suggested by the data, and therefore authors should report which analyses were prespecified. If subgroup analyses were undertaken, authors should report which subgroups were examined, why, whether they were prespecified, and how many were prespecified. Selective reporting of subgroup analyses could lead to bias.<sup>484</sup> When evaluating a subgroup, the question is not whether the subgroup demonstrates a statistically significant result but whether the subgroup treatment effects are significantly different from each other. To determine this, a test of interaction is helpful, although the power for such tests is typically low. If formal evaluations of interaction are undertaken (item 21d) they should be reported as the estimated difference in the intervention effect in each subgroup (with a CI), not just as P values.

In one survey,<sup>481</sup> 35 of 50 trial reports included subgroup analyses, of which only 42% used tests of interaction. It was often difficult to determine whether subgroup analyses had been specified in the protocol. In another survey of surgical trials published in high impact journals, 27 of 72 trials reported 54 subgroup analyses of which 91% were post hoc and only 6% of subgroup analyses used a test of interaction to assess whether a subgroup effect existed.<sup>485</sup>

#### **CONSORT 2025: Discussion**

**Item 29: Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence**

##### *Example*

"In this trial, which compared standard of care with monthly IPTp [intermittent preventive treatment in pregnancy] with sulfadoxine-pyrimethamine with monthly IPTp with dihydroartemisinin-piperaquine and monthly IPTp with dihydroartemisinin-piperaquine plus azithromycin, the use of dihydroartemisinin-piperaquine was associated with reductions in clinical malaria, malaria infection detected by microscopy during pregnancy, and placental malaria at the time of birth. However, despite these reductions in malaria, the risk of adverse pregnancy outcomes (composite primary endpoint) was significantly lower in the sulfadoxine-pyrimethamine group than in both dihydroartemisinin-piperaquine groups . . ."

"In a previous meta-analysis a modest, but nonsignificant, reduction of 17% in the number of adverse pregnancy outcomes favouring dihydroartemisinin-piperaquine versus sulfadoxine-pyrimethamine was reported . . . These findings led WHO [World Health Organization] to recommend larger definitive studies to determine whether IPTp with dihydroartemisinin-piperaquine improves adverse pregnancy outcomes compared with sulfadoxine-pyrimethamine in areas with high sulfadoxine-pyrimethamine resistance. Our results are consistent with those from a third trial in eastern Uganda, which showed that monthly IPTp with dihydroartemisinin-piperaquine did not decrease the number of adverse pregnancy outcomes among livebirths or reduce fetal loss compared with monthly IPTp with sulfadoxine-pyrimethamine . . . All of these trials, including ours, show that dihydroartemisinin-piperaquine is the superior antimalarial, resulting in 40–90% fewer malaria infections during pregnancy."<sup>196</sup>

#### *Explanation*

A good discussion section of a completed randomised trial should start with a brief summary of the trial results, balancing both benefits and harms of the intervention. The discussion sections of scientific reports are often filled with rhetoric supporting the authors' findings<sup>486</sup> and provide little measured argument of the pros and cons of the study and its results. Indeed, some discussion sections can be overly optimistic when discussing the trial findings (interpretation bias; spin).<sup>33 487-489</sup> Authors need to guard against such behaviours, as they diminish the rigour of the scientific effort and may result in a loss of trust by readers.

Readers will also want to know how the trial results relate to those of other randomised trials. This can best be achieved by including a published systematic review in the results or discussion section of the report.<sup>490 491</sup> This might be an easy ask for some authors as a systematic review might have been part

of the rationale for conducting the trial. However, for others, such synthesis may be impractical, and quoting and discussing any existing systematic reviews of similar trials may be more appropriate. One recent estimate suggests that nearly 80 systematic reviews are published daily.<sup>492</sup> Discussing trial findings in the context of results from any systematic reviews will help readers assess whether the results of the randomised trial are similar to those of other trials in the same topic area and whether participants are similar across studies. Reports of randomised trials have often not dealt adequately with these points.<sup>490</sup> Several methods to address the issue of setting new trial findings within the context of previous research have been proposed.<sup>491</sup> Where conducting a new updated systematic review is not practical, adding the new trial result to the previous systematic review is a much simpler alternative.

We recommend that at a minimum, the discussion should be as systematic and objective as possible and be based on a comprehensive search, rather than being limited to studies that support the results of the current trial.<sup>493 494</sup>

#### **Item 30: Trial limitations, addressing sources of potential bias, imprecision, generalisability, and, if relevant, multiplicity of analyses**

##### *Examples*

“The preponderance of male patients (85%) is a limitation of our study . . . We used bare-metal stents, since drug-eluting stents were not available until late during accrual. Although the latter factor may be perceived as a limitation, published data indicate no benefit (either short-term or long-term) with respect to death and myocardial infarction in patients with stable coronary artery disease who receive drug-eluting stents, as compared with those who receive bare-metal stents.”<sup>495</sup>

“Our study had several limitations. The early changes to the protocol to accommodate patients with a shorter injury history (but still not acute) to improve recruitment altered the characteristics of the study population. Overall, patients had less long-standing injury than was originally planned. Moreover, the study addressed a deliberately specific population of patients who continued to have ACL injury-related symptoms of instability and had not undergone any previous formal treatment. Another potential limitation is the proportion of patients who did not undergo surgical reconstruction, despite allocation to that group. The true benefit of surgical reconstruction could be somewhat greater than the ITT analysis suggests. The 18-month follow-up period ideally could have been longer but was constrained by various factors including funding. Notwithstanding, most patients had established their level of instability at this timepoint since being included in the trial. The trial design and analysis accounted for delayed surgery in both groups.”<sup>460</sup>

##### *Explanation*

An essential part of a good discussion section is summarising the limitations of the trial. Limitations

are frequently omitted from research reports<sup>496</sup>; identification and discussion of the weaknesses of a study have particular importance.

Some journals have attempted to remedy this problem by encouraging more structure to authors' discussion of their results.<sup>497 498</sup> For example, *The BMJ* recommends that authors structure the discussion section by presenting (1) a statement of the principal findings; (2) the strengths and weaknesses of the study; (3) the strengths and weaknesses in relation to other studies, discussing important differences in results; (4) the meaning of the study, its possible explanations and implications for clinicians and policymakers; and (5) any unanswered questions and future research.<sup>499</sup> We recommend that authors follow these sensible suggestions, perhaps also using suitable subheadings in the discussion section.

Authors should also discuss any imprecision of the results. Imprecision may arise in connection with several aspects of a study, including measurement of a primary outcome (item 14) or diagnosis (item 12a). Perhaps the scale used was validated on an adult population but used in a paediatric one, or the assessor was not trained in how to administer the instrument.

The difference between statistical significance and clinical importance should always be borne in mind. Authors should particularly avoid the common error of interpreting a non-significant result as indicating equivalence of interventions. The CI (item 26) provides valuable insight into whether the trial result is compatible with a clinically important effect, regardless of the P value.<sup>500</sup>

Authors should exercise special care when evaluating the results of trials with multiple comparisons. Such multiplicity arises from several interventions, outcome measures, time points, subgroup analyses, and other factors. In such circumstances, some statistically significant findings are likely to result from chance alone.

Authors should also consider the extent to which the results of a study can be generalised to other circumstances; also known as external validity. For example, can the results be generalised to an individual participant or groups that differ from those enrolled in the trial with regard to age, sex, severity of disease, and comorbid conditions? Are the results applicable to other drugs within a class of similar drugs, to a different dose, timing, and route of administration? Can similar results be expected in different healthcare settings? What about the effect on related outcomes that were not assessed in the trial, and the importance of length of follow-up and duration of treatment, especially with respect to harms?<sup>501</sup> Internal validity, the extent to which the design and conduct of the trial eliminates the possibility of bias, is a prerequisite for external validity: the results of a flawed trial are invalid and the question of its external validity becomes irrelevant. External validity is a matter of judgement and depends on the characteristics of the participants included in the trial, the trial setting, the treatment regimens tested, and the outcomes assessed.

## Discussion

It is critical that every randomised trial has a complete and transparent report of the trial methods and findings, to enable readers to judge its reliability and validity or to extract information for inclusion systematic reviews. The aim of this explanation and elaboration document is to assist authors in using CONSORT 2025 and explain in general terms the importance of adequate reporting of trials, along with providing examples of good reporting. The CONSORT 2025 statement can also help guide peer reviewers and journal editors in their evaluation of manuscripts. Indeed, we encourage peer reviewers and editors to use the CONSORT 2025 checklist to assess whether authors have reported on these items. Such assessments are likely to improve the clarity and transparency of published trials.<sup>3 502</sup>

CONSORT is a minimum set of essential items that should be included in reports of randomised trials, published in the form of the CONSORT 2025 statement,<sup>24</sup> this explanation and elaboration document, and a methods paper describing the development process.<sup>17</sup>

Key strengths of CONSORT 2025 include its systematic and transparent development methods, the participation of a wide range of key stakeholders (including via a large international Delphi survey), the use of empirical evidence to support its recommendations, and the availability of detailed guidance including recent examples of good reporting from published trials. Developing CONSORT 2025 and SPIRIT 2025 together was also an opportunity to align both checklists and to provide users with consistent guidance in the reporting of trial design, conduct, and analysis, from trial protocol to final publication.

A challenge with CONSORT 2025, as with any reporting guideline, is to achieve a reasonable balance between brevity versus comprehensiveness. The CONSORT 2025 checklist is intended to specify the minimum items to address in reports of randomised trials, and we have tried to keep the checklist as short as possible. We have developed an expanded checklist (appendix 1) incorporating the core content of the explanation and elaboration paper in short format, which authors may be helpful when writing their trial publication. This CONSORT 2025 explanation and elaboration paper, including its comprehensive list of references, serves to fulfil the needs of those seeking more detail and context. A joint SPIRIT-CONSORT website (<https://www.consort-spirit.org/>) has been established to provide more information about the CONSORT and SPIRIT statements, including additional resources and training materials aimed at researchers, research trainees, journal editors, and peer reviewers. It also includes resources aimed at patients and the public that explain the importance of transparent reporting of randomised trials and their importance in the delivery of evidence based healthcare.

The CONSORT statement is an evolving document, which requires a dynamic process of continual assessment, refinement, and, if necessary, change, which is why we have updated the checklist and this

explanation and elaboration article. As new evidence and critical comments accumulate, we will evaluate the need for future updates.

## AUTHOR AFFILIATIONS

- <sup>1</sup>Oxford Clinical Trials Research Unit, Centre for Statistics in Medicine, University of Oxford, Oxford OX3 7LD, UK
- <sup>2</sup>Department of Medicine, Women's College Research Institute, University of Toronto, Toronto, ON, Canada
- <sup>3</sup>UK EQUATOR Centre, Centre for Statistics in Medicine, University of Oxford, Oxford, UK
- <sup>4</sup>Centre for Evidence-Based Medicine Odense and Cochrane Denmark, Department of Clinical Research, University of Southern Denmark, Odense, Denmark
- <sup>5</sup>Open Patient data Explorative Network, Odense University Hospital, Odense, Denmark
- <sup>6</sup>Centre for Journalology, Clinical Epidemiology Programme, Ottawa Hospital Research Institute, Ottawa, ON, Canada
- <sup>7</sup>Department of Obstetrics and Gynecology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
- <sup>8</sup>Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India
- <sup>9</sup>Office of Science Dissemination, Centers for Disease Control and Prevention, Atlanta, GA, USA
- <sup>10</sup>Department of Biostatistics and Epidemiology, School of Public Health, Center for Pharmacoepidemiology and Treatment Science, Rutgers University, New Brunswick, NJ, USA
- <sup>11</sup>*JAMA Network Open*, Chicago, IL, USA
- <sup>12</sup>Centre for Health Research and Development, Society for Applied Studies, New Delhi, India
- <sup>13</sup>Child Health Evaluation Services, The Hospital for Sick Children Research Institute, Toronto, ON, Canada
- <sup>14</sup>Department of Psychiatry, University of Toronto, Toronto, ON, Canada
- <sup>15</sup>Aberdeen Centre for Evaluation, University of Aberdeen, Aberdeen, UK
- <sup>16</sup>Project PINK BLUE - Health & Psychological Trust Centre, Utako, Abuja, Nigeria
- <sup>17</sup>Department of Sociology and Gerontology, Miami University, OH, USA
- <sup>18</sup>Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK
- <sup>19</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
- <sup>20</sup>Ottawa Hospital Research Institute, Ottawa, ON, Canada
- <sup>21</sup>Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
- <sup>22</sup>Department of Epidemiology and Population Health, Stanford University, Palo Alto, CA, USA
- <sup>23</sup>Institute for Evidence-Based Healthcare, Faculty of Health Sciences and Medicine, Bond University, University Drive, Robina, QLD, Australia
- <sup>24</sup>Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Statistics, and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA
- <sup>25</sup>MRC Clinical Trials Unit at University College London, London, UK
- <sup>26</sup>University College London, UCL Great Ormond Street Institute of Child Health, London, UK
- <sup>27</sup>NIHR Exeter Biomedical Research Centre, Faculty of Health and Life Sciences, University of Exeter, Exeter, UK
- <sup>28</sup>Edinburgh Clinical Trials Unit, Usher Institute-University of Edinburgh, Edinburgh BioQuarter, Edinburgh, UK
- <sup>29</sup>*The BMJ*, BMA House, London, UK
- <sup>30</sup>Harvard Medical School, Boston, MA, USA
- <sup>31</sup>Université Paris Cité, Inserm, INRAE, Centre de Recherche Épidémiologie et Statistiques, Université Paris Cité, Paris, France
- <sup>32</sup>Clinical Trials Ontario, MaRS Centre, Toronto, ON, Canada
- <sup>33</sup>Duke Clinical Research Institute, Duke University Medical Center, Durham, NC, USA

<sup>34</sup>Department of Emergency Medicine, University of California, Los Angeles, CA, USA

<sup>35</sup>South African Medical Research Council, Cape Town, South Africa

<sup>36</sup>Warwick Research in Nursing, Warwick Medical School, University of Warwick, Coventry, UK

<sup>37</sup>MRC/CSO Social and Public Health Sciences Unit & Robertson Centre for Biostatistics, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

<sup>38</sup>Department of Health Research Methods Evidence and Impact, McMaster University, Hamilton, ON, Canada

<sup>39</sup>St Joseph's Healthcare Hamilton, Hamilton, ON, Canada

<sup>40</sup>York Trials Unit, Department of Health Sciences, University of York, York, UK

<sup>41</sup>Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

<sup>42</sup>Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Centre for Research in Epidemiology and Statistics (CRESS), Paris, France

<sup>43</sup>Centre d'Epidémiologie Clinique, Hôpital Hôtel Dieu, AP-HP, Paris, France

**SPIRIT-CONSORT executive group:** Sally Hopewell, University of Oxford, UK; An-Wen Chan, University of Toronto, Canada; Asbjørn Hróbjartsson, University of Southern Denmark, Odense, Denmark; David Moher, University of Ottawa, Canada; Kenneth Schulz, University of North Carolina at Chapel Hill, USA; Isabelle Boutron, Université Paris Cité, France. Gary Collins and Ruth Tunn, both of University of Oxford, UK, were also involved in leading the SPIRIT and CONSORT 2025 update.

**SPIRIT-CONSORT 2025 consensus meeting participants:** Rakesh Aggarwal, Jawaharlal Institute of Postgraduate Medical Education and Research, India; Michael Berkwits, Office of the Centers for Disease Control and Prevention, USA (formally JAMA and the JAMA Network at time of consensus meeting); Jesse A Berlin, Rutgers University/JAMA Network Open USA; Nita Bhandari, Society for Applied Studies, India; Nancy J Butcher, The Hospital for Sick Children, Canada; Marion K Campbell, University of Aberdeen, UK; Runcie CW Chidebe, Project PINK BLUE, Nigeria/Miami University, Ohio, USA; Diana Elbourne, London School of Hygiene and Tropical Medicine, UK; Andrew J Farmer, University of Oxford, UK; Dean A Ferguson, Ottawa Hospital Research Institute, Canada; Robert M Golub, Northwestern University, USA; Steven N Goodman, Stanford University, USA; Tammy C Hoffmann, Bond University, Australia; John PA Ioannidis, Stanford University, USA; Brennan C Kahan, University College London, UK; Rachel L Knowles, University College London, UK; Sarah E Lamb, University of Exeter, UK; Steff Lewis, University of Edinburgh, UK; Elizabeth Loder, *The BMJ*, UK; Martin Offringa, Hospital for Sick Children Research Institute, Canada; Dawn P Richards, Clinical Trials Ontario, Canada; Frank W Rockhold, Duke University, USA; David L Schriger, University of California, USA; Nandi L Siegfried, South African Medical Research Council, South Africa; Sophie Staniszewska, University of Warwick, UK; Rod S Taylor, University of Glasgow, UK; Lehana Thabane, McMaster University/St Joseph's Healthcare, Canada; David Torgerson, University of York, UK; Sunita Vohra, University of Alberta, Canada; and Ian R White, University College London, UK.

We dedicate CONSORT 2025 to the late Doug Altman who was instrumental in the development of the SPIRIT and CONSORT statements and whose files and correspondence contributed to this update following his death. We gratefully acknowledge the contributions of all those who participated in the SPIRIT-CONSORT 2025 update Delphi survey. We also acknowledge Camilla Hansen Nejstgaard for conducting the scoping review of the literature to identify suggested changes to CONSORT 2010 and for comparing CONSORT 2025 with RoB2, Lasse Østengård for developing the SCEB (SPIRIT-CONSORT Evidence Bibliographic) database of empirical evidence to support the development of CONSORT, Jen de Bayer and Patricia Logullo for their involvement during the consensus meeting, and Lina El Ghosn for initial drafting of the expanded CONSORT checklist.

**Contributors:** SH, IB, A-WC, AH, KFS, and DM were responsible for the concept; SH, IB, A-WC, GSC, AH, KFS, and DM were responsible for the funding. SH, IB, A-WC, GSC, AH, KFS, RT, and DM drafted specific sections of the manuscript. All authors then critically revised the manuscript for important intellectual content and approved the final version. SH is the guarantor of this work. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Funding:** The 2025 update of SPIRIT and CONSORT was funded by the MRC-NIHR: Better Methods, Better Research (MR/W020483/1). The funder reviewed the design of the study but had no role in the collection, analysis, or interpretation of the data; in the writing of the report; or in the decision to submit the article for publication. The views expressed are those of the author(s) and not necessarily those of the NIHR, the MRC, or the Department of Health and Social Care.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form at <https://www.icmje.org/conflicts-of-interest/> and declare: support from MRC-NIHR for the submitted work. SH, IB, A-WC, AH, KFS, and DM are members of the SPIRIT-CONSORT executive group. SH, IB, A-WC, AH, KFS, GSC, DM, MKC, NJB, MO, RST, and SV are involved in the development, update, implementation, and dissemination of several reporting guidelines. GSC is the director of the UK EQUATOR Centre, a statistical editor for *The BMJ* and NIHR Senior Investigator, DM is the director of the Canadian EQUATOR Centre, and member of *The BMJ*'s regional advisory board for North America, IB is deputy director and PR is director of the French EQUATOR Centre, TCH is director of the Australasian EQUATOR Centre, JPAI is director of the US EQUATOR Centre. RA is president of the World Association of Medical Editors. MKC is chair of the MRC-NIHR: Better Methods Better Research funding panel. RCWC is executive director of Project PINK-BLUE, which receives funding from Roche-Product. AF is director of the UK National Institute for Health and Care Research Health Technology Assessment Programme. DPR is a full time employee of FiveO2 Laboratories, which under contract to Clinical Trials Ontario provides services related to patient and public engagement; and is the volunteer vice president of the Canadian Arthritis Patient Alliance, which receives funding through independent grants from pharmaceutical companies. IRW was supported by the MRC Programmes MCUU00004/07 and MCUU00004/09. DLS is *JAMA* associate editor and receives editing stipends from *JAMA* and *Annals of Emergency Medicine*.

**Ethical approval:** Ethical approval was granted by the Central University Research Ethics Committee, University of Oxford (R76421/RE001). All Delphi participants provided informed consent to participate.

**Patient and public involvement:** The CONSORT 2025 checklist items and the explanations here were developed using input from an international Delphi survey and consensus meeting. The Delphi survey was advertised via established patient and public involvement (PPI) networks, and 17 respondents self-identified as a "patient or public representative" and completed the Delphi survey. In addition, three of the participants in the expert consensus meeting were patient or public representatives who were leaders in advancing PPI.

**Dissemination to participants and related patient and public communities:** CONSORT 2025 will be disseminated via a new website, [consort-spirit.org](http://consort-spirit.org), which will include materials designed for patients and the public.

**Provenance and peer review:** Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>.

- Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996;313:570-1. doi:10.1136/bmj.313.7057.570
- Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383:267-76. doi:10.1016/S0140-6736(13)62228-X
- Turner L, Shamseer L, Altman DG, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev* 2012;11:MR000030. doi:10.1002/14651858.MR000030.pub2
- Savović J, Jones HE, Altman DG, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med* 2012;157:429-38. doi:10.7326/0003-4819-157-6-201209180-00537
- Goldacre B, Drysdale H, Powell-Smith A, et al. The COMParE Trials Project. <https://www.compare-trials.org/> 2016.
- Andrew E, Anis A, Chalmers T, et al, The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. *JAMA* 1994;272:1926-31. doi:10.1001/jama.1994.035204054041

- 7 The Asilomar Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. Checklist of information for inclusion in reports of clinical trials. *Ann Intern Med* 1996;124:741-3. doi:10.7326/0003-4819-124-8-199604150-00007
- 8 Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637-9. doi:10.1001/jama.1996.03540080059030
- 9 Moher D, Schulz KF, Altman D, CONSORT Group (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285:1987-91. doi:10.1001/jama.285.15.1987
- 10 Altman DG, Schulz KF, Moher D, et al, CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663-94. doi:10.7326/0003-4819-134-8-200104170-00012
- 11 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
- 12 Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869. doi:10.1136/bmj.c869
- 13 Chan A-W, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200-7. doi:10.7326/0003-4819-158-3-201302050-00583
- 14 Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586. doi:10.1136/bmj.e7586
- 15 Hopewell S, Boutron I, Chan A-W, et al. An update to SPIRIT and CONSORT reporting guidelines to enhance transparency in randomized trials. *Nat Med* 2022;28:1740-3. doi:10.1038/s41591-022-01989-8
- 16 Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010;7:e1000217. doi:10.1371/journal.pmed.1000217
- 17 Tunn R, Boutron I, Chan AW, et al. Methods used to develop the SPIRIT 2024 and CONSORT 2024 Statements. *J Clin Epidemiol* 2024;169:111309. doi:10.1016/j.jclinepi.2024.111309
- 18 Nejstgaard CH, Boutron I, Chan AW, et al. A scoping review identifies multiple comments suggesting modifications to SPIRIT 2013 and CONSORT 2010. *J Clin Epidemiol* 2023;155:48-63. doi:10.1016/j.jclinepi.2023.01.003
- 19 Østenggaard L, Barrientos A, Boutron I, et al. Development of a topic-specific bibliographic database supporting the updates of SPIRIT 2013 and CONSORT 2010. *Cochrane Evidence Synthesis and Methods* 2024;2:e12057. doi:10.1002/cesm.12057
- 20 Junqueira DR, Zorzela L, Golder S, et al, CONSORT Harms Group. CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *BMJ* 2023;381:e073725. doi:10.1136/bmj-2022-073725
- 21 Butcher NJ, Monsour A, Mew Ej, et al. Guidelines for reporting outcomes in trial reports: The CONSORT-Outcomes 2022 extension. *JAMA* 2022;328:2252-64. doi:10.1001/jama.2022.21022
- 22 Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P, CONSORT NPT Group. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Ann Intern Med* 2017;167:40-7. doi:10.7326/M17-0046
- 23 Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:g1687. doi:10.1136/bmj.g1687
- 24 Hopewell S, Chan A-W, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. *BMJ* 2025;389:e081123. doi:10.1136/bmj-2024-081123.
- 25 Bouzat P, Charbit J, Abback P-S, et al, PROCOAG Study Group. Efficacy and safety of early administration of 4-factor prothrombin complex concentrate in patients with trauma at risk of massive transfusion: the PROCOAG randomized clinical trial. *JAMA* 2023;329:1367-75. doi:10.1001/jama.2023.4080
- 26 Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Eval Health Prof* 2002;25:38-64.
- 27 PLoS Medicine Editors. The impact of open access upon public health. *PLoS Med* 2006;3:e252. doi:10.1371/journal.pmed.0030252
- 28 Harbour AM, Knecht LS, Humphreys BL. Structured abstracts in MEDLINE, 1989-1991. *Bull Med Libr Assoc* 1995;83:190-5.
- 29 Harris AHS, Standard S, Brunning JL, et al. The accuracy of abstracts in psychology journals. *J Psychol* 2002;136:141-8. doi:10.1080/00223980209604145
- 30 Pitkin RM, Branagan MA, Burmeister LF. Accuracy of data in abstracts of published research articles. *JAMA* 1999;281:1110-1. doi:10.1001/jama.281.12.1110
- 31 Ward LG, Kendrach MG, Price SO. Accuracy of abstracts for original research articles in pharmacy journals. *Ann Pharmacother* 2004;38:1173-7. doi:10.1345/aph.1D416
- 32 Gøtzsche PC. Believability of relative risks and odds ratios in abstracts: cross sectional study. *BMJ* 2006;333:231-4. doi:10.1136/bmj.38895.41045179
- 33 Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA* 2010;303:2058-64. doi:10.1001/jama.2010.651
- 34 Chiu K, Grundy Q, Bero L. 'Spin' in published biomedical literature: A methodological systematic review. *PLoS Biol* 2017;15:e2002173. doi:10.1371/journal.pbio.2002173
- 35 Spungen H, Burton J, Schenkel S, Schriger DL. Completeness and Spin of medRxiv Preprint and Associated Published Abstracts of COVID-19 Randomized Clinical Trials. *JAMA* 2023;329:1310-2. doi:10.1001/jama.2023.1784
- 36 Ioannidis JPA, Evans SJW, Gøtzsche PC, et al, CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;141:781-8. doi:10.7326/0003-4819-141-10-200411160-00009
- 37 Ioannidis JPA, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA* 2001;285:437-43. doi:10.1001/jama.285.4.437
- 38 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:e20. doi:10.1371/journal.pmed.0050020
- 39 Chhapola V, Tiwari S, Brar R, Kanwal SK. Reporting quality of trial abstracts-improved yet suboptimal: A systematic review and meta-analysis. *J Evid Based Med* 2018;11:89-94. doi:10.1111/jebm.12294
- 40 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:e20. doi:10.1371/journal.pmed.0050020
- 41 Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More informative abstracts revisited. *Ann Intern Med* 1990;113:69-76. doi:10.7326/0003-4819-113-1-69
- 42 Taddio A, Pain T, Fassos FF, Boon H, Ilersich AL, Einarsen TR. Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal, the Canadian Medical Association Journal and the Journal of the American Medical Association. *CMAJ* 1994;150:1611-5.
- 43 Wager E, Middleton P. Technical editing of research reports in biomedical journals. *Cochrane Database Syst Rev* 2008;2008:MR000002. doi:10.1002/14651858.MR000002.pub3
- 44 Hartley J, Sydes M, Bl Burton A. Obtaining information accurately and quickly: are structured abstracts more efficient? *Inf Sci* 1996;22:349-56. doi:10.1177/016555159602200503.
- 45 UNESCO. UNESCO Recommendation on Open Science. <https://unesdoc.unesco.org/ark:/48223/pf000379949>. 2021.
- 46 Ganjavi C, Eppler MB, Pekcan A, et al. Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. *BMJ* 2024;384:e077192. doi:10.1136/bmj-2023-077192
- 47 Nazari M, Shabani R, Hassanzadeh-Rad A, Esfandiari MA, Dallili S. Effect of concurrent resistance-aerobic training on inflammatory factors and growth hormones in children with type 1 diabetes: a randomized controlled clinical trial. *Trials* 2023;24:519. doi:10.1186/s13063-023-07553-0
- 48 Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997;9(Suppl):15-21.
- 49 Alayache M, Cobey KD, Ng JY, et al. Evaluating prospective study registration and result reporting of trials conducted in Canada from 2009 to 2019. *Facets* 2023;8:1-10. doi:10.1139/facets-2022-0208.
- 50 Riedel N, Wieschowski S, Bruckner T, et al. Results dissemination from completed clinical trials conducted at German university medical centers remained delayed and incomplete. The 2014-2017 cohort. *J Clin Epidemiol* 2022;144:1-7. doi:10.1016/j.jclinepi.2021.12.012
- 51 Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65. doi:10.1001/jama.291.20.2457
- 52 Al-Marzouki S, Roberts I, Evans S, Marshall T. Selective reporting in clinical trials: analysis of trial protocols accepted by The Lancet. *Lancet* 2008;372:201. doi:10.1016/S0140-6736(08)61060-0
- 53 Williamson PR, Gamble C. Identification and impact of outcome selection bias in meta-analysis. *Stat Med* 2005;24:1547-61. doi:10.1002/sim.2025

- 54 Simes RJ. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol* 1986;4:1529-41. doi:10.1200/JCO.1986.4.10.1529
- 55 Chalmers I. From optimism to disillusion about commitment to transparency in the medico-industrial complex. *J R Soc Med* 2006;99:337-41. doi:10.1177/014107680609900715
- 56 Tonks A. A clinical trials register for Europe. *BMJ* 2002;325:1314-5. doi:10.1136/bmj.325.7376.1314
- 57 Dickersin K, Rennie D. Registering clinical trials. *JAMA* 2003;290:516-23. doi:10.1001/jama.290.4.516
- 58 Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, Boddington E. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004;363:1341-5. doi:10.1016/S0140-6736(04)16043-1
- 59 World Health Organization. *International standards for clinical trial registries: the registration of all interventional trials is a scientific, ethical and moral responsibility*. 3rd ed. World Health Organization, 2018.
- 60 De Angelis CD, Drazen JM, Frizelle FA, et al. International Committee of Medical Journal Editors. Is this clinical trial fully registered? A statement from the International Committee of Medical Journal Editors. *Lancet* 2005;365:1827-9. doi:10.1016/S0140-6736(05)66588-9
- 61 Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DAB. Issues in the registration of clinical trials. *JAMA* 2007;297:2112-20. doi:10.1001/jama.297.19.2112
- 62 Nosek BA, Alter G, Banks GC, et al. SCIENTIFIC STANDARDS. Promoting an open research culture. *Science* 2015;348:1422-5. doi:10.1126/science.aab2374
- 63 Shamseer L, Hopewell S, Altman DG, Moher D, Schulz KF. Update on the endorsement of CONSORT by high impact factor journals: a survey of journal "Instructions to Authors" in 2014. *Trials* 2016;17:301. doi:10.1186/s13063-016-1408-z
- 64 Lombard N, Gasmi A, Sulpice L, Boudjema K, Naudet F, Bergeat D. Research transparency promotion by surgical journals publishing randomised controlled trials: a survey. *Trials* 2020;21:824. doi:10.1186/s13063-020-04756-7
- 65 Trinquart L, Dunn AG, Bourgeois FT. Registration of published randomized trials: a systematic review and meta-analysis. *BMC Med* 2018;16:173. doi:10.1186/s12916-018-1168-6
- 66 Chan A-W, Pello A, Kitchen J, et al. Association of Trial Registration With Reporting of Primary Outcomes in Protocols and Publications. *JAMA* 2017;318:1709-11. doi:10.1001/jama.2017.13001
- 67 Chan AW, Song F, Vickers A, et al. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;383:257-66. doi:10.1016/S0140-6736(13)62296-5
- 68 DeVito NJ, Bacon S, Goldacre B. Compliance with legal requirement to report clinical trial results on ClinicalTrials.gov: a cohort study. *Lancet* 2020;395:361-9. doi:10.1016/S0140-6736(19)33220-9
- 69 DeVito NJ, Morley J, Smith JA, Drysdale H, Goldacre B, Heneghan C. Availability of results of clinical trials registered on EU Clinical Trials Register: cross sectional audit study. *BMJ Med* 2024;3:e000738. doi:10.1136/bmjjmed-2023-000738
- 70 Pellat A, Boutron I, Ravaud P. Availability of Results of Trials Studying Pancreatic Adenocarcinoma over the Past 10 Years. *Oncologist* 2022;27:e849-55. doi:10.1093/oncology/oyac156
- 71 Moorthy VS, Karam G, Vannice KS, Kiely MP. Rationale for WHO's new position calling for prompt reporting and public disclosure of interventional clinical trial results. *PLoS Med* 2015;12:e1001819. doi:10.1371/journal.pmed.1001819
- 72 Food and Drug Administration Amendments Act of 2007. Public Law 110-85. <https://www.govinfo.gov/app/details/PLAW-110publ85d> [accessed 17 September 2024]. 2007.
- 73 European Commission. Commission Guideline — Guidance on posting and publication of result-related information on clinical trials in relation to the implementation of Article 57(2) of Regulation (EC) No 726/2004 and Article 41(2) of Regulation (EC) No 1901/2006. *Official J Eur Union*. 2012;55:7-10.
- 74 Gross O, Tönshoff B, Weber LT, et al. German Pediatric Nephrology (GPN) Study Group and EARLY PRO-TECT Alport Investigators. A multicenter, randomized, placebo-controlled, double-blind phase 3 trial with open-arm comparison indicates safety and efficacy of nephroprotective therapy with ramipril in children with Alport's syndrome. *Kidney Int* 2020;97:1275-86. doi:10.1016/j.kint.2019.12.015
- 75 Goldacre B, Drysdale H, Dale A, et al. COMParE: a prospective cohort study correcting and monitoring 58 misreported trials in real time. *Trials* 2019;20:118. doi:10.1186/s13063-019-3173-2
- 76 Goldacre B, Drysdale H, Marston C, et al. COMPAR-E: Qualitative analysis of researchers' responses to critical correspondence on a cohort of 58 misreported trials. *Trials* 2019;20:124. doi:10.1186/s13063-019-3172-3
- 77 Chan A-W, Hróbjartsson A. Promoting public access to clinical trial protocols: challenges and recommendations. *Trials* 2018;19:116. doi:10.1186/s13063-018-2510-1
- 78 Chan A-W, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. *BMJ* 2025;389:e081477. doi:10.1136/bmj-2024-081477
- 79 Gamble C, Krishan A, Stocken D, et al. Guidelines for the content of statistical analysis plans in clinical trials. *JAMA* 2017;318:2337-43. doi:10.1001/jama.2017.18556
- 80 Rangan A, Brealey SD, Keding A, et al. UK FROST Study Group. Management of adults with primary frozen shoulder in secondary care (UK FROST): a multicentre, pragmatic, three-arm, superiority randomised clinical trial. *Lancet* 2020;396:977-89. doi:10.1016/S0140-6736(20)31965-6
- 81 Roaldsen MB, Eltoft A, Wilsgaard T, et al. TWIST Investigators. Safety and efficacy of tenecteplase in patients with wake-up stroke assessed by non-contrast CT (TWIST): a multicentre, open-label, randomised controlled trial. *Lancet Neurol* 2023;22:117-26. doi:10.1016/S1474-4422(22)00484-7
- 82 Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B. The citation advantage of linking publications to research data. *PLOS One* 2020;15:e0230416. doi:10.1371/journal.pone.0230416
- 83 Stewart LA, Clarke M, Rovers M, et al. PRISMA-IPD Development Group. Preferred reporting items for a systematic review and meta-analysis of individual participant data: the PRISMA-IPD statement. *JAMA* 2015;313:1657-65. doi:10.1001/jama.2015.3656
- 84 Naudet F, Sakarovitch C, Janiaud P, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in *The BMJ* and *PLOS Medicine*. *BMJ* 2018;360:k400. doi:10.1136/bmj.k400
- 85 Bergeat D, Lombard N, Gasmi A, Le Floch B, Naudet F. Data sharing and reanalyses among randomized clinical trials published in surgical journals before and after adoption of a data availability and reproducibility policy. *JAMA Netw Open* 2022;5:e2215209. doi:10.1001/jamanetworkopen.2022.15209
- 86 Rowhani-Farid A, Grewal M, Solar S, et al. Clinical trial data sharing: a cross-sectional study of outcomes associated with two U.S. National Institutes of Health models. *Sci Data* 2023;10:529. doi:10.1038/s41597-023-02436-0
- 87 Esmail LC, Kapp P, Assi R, et al. Sharing of individual patient-level data by trialists of randomized clinical trials of pharmacological treatments for COVID-19. *JAMA* 2023;329:1695-7. doi:10.1001/jama.2023.4590
- 88 Ohmann C, Moher D, Siebert M, Motschall E, Naudet F. Status, use and impact of sharing individual participant data from clinical trials: a scoping review. *BMJ Open* 2021;11:e049228. doi:10.1136/bmjopen-2021-049228
- 89 Larregue J, Vincent-Lamarre P, Lebaron F, Larivière V. COVID-19: Where is the data? LSE Impact Blog, 30 novembre 2020. <https://unesco.ebsi.umontreal.ca/en/publications/covid-19-where-is-the-data/>
- 90 Ioannidis JPA. Hundreds of thousands of zombie randomised trials circulate among us. *Anaesthesia* 2021;76:444-7. doi:10.1111/anae.15297
- 91 Carlisle JB. False individual patient data and zombie randomised controlled trials submitted to *Anaesthesia*. *Anaesthesia* 2021;76:472-9. doi:10.1111/anae.15263
- 92 Pellen C, Le Louarn A, Spurrier-Bernard G, et al. Ten (not so) simple rules for clinical trial data-sharing. *PLoS Comput Biol* 2023;19:e1010879. doi:10.1371/journal.pcbi.1010879
- 93 Tudor Smith C, Hopkins C, Sydes M, et al. Good Practice Principles for Sharing Individual Participant Data from Publicly Funded Clinical Trials. <https://www.methodologyhubs.mrc.ac.uk/files/7114/3682/3831/Datasharingguidance2015.pdf> [accessed 15 April 2024]. 2015.
- 94 The White House. OSTP Issues Guidance to Make Federally Funded Research Freely Available Without Delay. <https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/> [accessed 17 September 2024]. 2022.
- 95 National Institutes of Health. Data management and sharing policy. <https://sharing.nih.gov/data-management-and-sharing-policy/data-management#:~:text=NIH%20encourages%20data%20management%20and,repurposing%20datasets%20for%20secondary%20research.> [Accessed 29 Jan 2024]
- 96 Siebert M, Ioannidis JPA. Lifting of Embargoes to Data Sharing in Clinical Trials Published in Top Medical Journals. *JAMA* 2024;331:354-5. doi:10.1001/jama.2023.25394
- 97 Gregson S, Adamson S, Papaya S, et al. Impact and process evaluation of integrated community and clinic-based HIV-1 control: a cluster-randomised trial in eastern Zimbabwe. *PLoS Med* 2007;4:e102. doi:10.1371/journal.pmed.0040102
- 98 Ha SY, Oh DY, Yañez P, et al. KEYNOTE-859 investigators. Pembrolizumab plus chemotherapy versus placebo plus chemotherapy for HER2-negative advanced gastric cancer (KEYNOTE-859): a multicentre, randomised, double-blind, phase 3 trial. *Lancet Oncol* 2023;24:1181-95. doi:10.1016/S1470-4045(23)00515-6

- 99 Speich B, von Niederhäusern B, Schur N, et al, Making Randomized Trials Affordable (MARTA) Group. Systematic review on costs and resource use of randomized clinical trials shows a lack of transparent and comprehensive data. *J Clin Epidemiol* 2018;96:1-11. doi:10.1016/j.jclinepi.2017.12.018
- 100 Martin L, Hutchens M, Hawkins C, Radnov A. How much do clinical trials cost? *Nat Rev Drug Discov* 2017;16:381-2. doi:10.1038/nrd.2017.70
- 101 Raftery J, Young A, Stanton L, et al. Clinical trial metadata: defining and extracting metadata on the design, conduct, results and costs of 125 randomised clinical trials funded by the National Institute for Health Research Health Technology Assessment programme. *Health Technol Assess* 2015;19:1-138. doi:10.3310/hta19110
- 102 Rasmussen K, Bero L, Redberg R, Götzsche PC, Lundh A. Collaboration between academics and industry in clinical trials: cross sectional study of publications and survey of lead academic authors. *BMJ* 2018;363:k3654. doi:10.1136/bmj.k3654
- 103 Siena LM, Papamanolis L, Siebert MJ, Bellomo RK, Ioannidis JPA. Industry Involvement and Transparency in the Most Cited Clinical Trials, 2019-2022. *JAMA Netw Open* 2023;6:e2343425. doi:10.1001/jamanetworkopen.2023.43425
- 104 Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev* 2017;2:MR000033.
- 105 Nejstgaard CH, Laursen DRT, Lundh A, Hróbjartsson A. Commercial funding and estimated intervention effects in randomized clinical trials: Systematic review of meta-epidemiological studies. *Res Synth Methods* 2023;14:144-55. doi:10.1002/rsm.1611
- 106 Østengaard L, Lundh A, Tjørnhøj-Thomsen T, et al. Influence and management of conflicts of interest in randomised clinical trials: qualitative interview study. *BMJ* 2020;371:m3764. doi:10.1136/bmj.m3764
- 107 Hakoum MB, Jouni N, Abou-Jaoude EA, et al. Characteristics of funding of clinical trials: cross-sectional survey and proposed guidance. *BMJ Open* 2017;7:e015997. doi:10.1136/bmjopen-2017-015997
- 108 Institute of Medicine. *Conflict of interest in medical research, education, and practice*. National Academies Press, 2009.
- 109 Hakoum MB, Jouni N, Abou-Jaoude EA, et al. Authors of clinical trials reported individual and financial conflicts of interest more frequently than institutional and nonfinancial ones: a methodological survey. *J Clin Epidemiol* 2017;87:78-86. doi:10.1016/j.jclinepi.2017.04.002
- 110 Lenzer J. When is a point of view a conflict of interest? *BMJ* 2016;355:i6194. doi:10.1136/bmj.i6194
- 111 Ahn R, Woodbridge A, Abraham A, et al. Financial ties of principal investigators and randomized controlled trial outcomes: cross sectional study. *BMJ* 2017;356:i6770. doi:10.1136/bmj.i6770
- 112 El-Rayess H, Khamis AM, Haddad S, et al. Assessing concordance of financial conflicts of interest disclosures with payments' databases: a systematic survey of the health literature. *J Clin Epidemiol* 2020;127:19-28. doi:10.1016/j.jclinepi.2020.06.040
- 113 Rasmussen K, Schroll J, Götzsche PC, Lundh A. Under-reporting of conflicts of interest among trialists: a cross-sectional study. *J R Soc Med* 2015;108:101-7. doi:10.1177/014107614557878
- 114 Hopewell S, Keene DJ, Marian IR, et al, GRASP Trial Group. Progressive exercise compared with best practice advice, with or without corticosteroid injection, for the treatment of patients with rotator cuff disorders (GRASP): a multicentre, pragmatic, 2 × 2 factorial, randomised controlled trial. *Lancet* 2021;398:416-28. doi:10.1016/S0140-6736(21)00846-1
- 115 Dawson L, Zarin DA, Emanuel EJ, Friedman LM, Chaudhari B, Goodman SN. Considering usual medical care in clinical trial design. *PLoS Med* 2009;6:e1000111. doi:10.1371/journal.pmed.1000111
- 116 van Luijn JC, van Loenen AC, Gribnau FW, Leufkens HG. Choice of comparator in active control trials of new drugs. *Ann Pharmacother* 2008;42:1605-12. doi:10.1345/aph.1L115
- 117 World Medical Association Declaration of Helsinki, ethical principle for medical research involving human participants. 75th WMA General Assembly, Finland, 2024. <https://www.wma.net/policies-post/wma-declaration-of-helsinki/> (accessed 10 February 2025).
- 118 Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992;327:248-54. doi:10.1056/NEJM199207233270406
- 119 Ferguson D, Glass KC, Hutton B, Shapiro S. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? *Clin Trials* 2005;2:218-29, discussion 229-32. doi:10.1191/1740774505cn085oa
- 120 Clarke M, Brice A, Chalmers I. Accumulating research: a systematic account of how cumulative meta-analyses would have provided knowledge, improved health, reduced harm and saved resources. *PLoS One* 2014;9:e102670. doi:10.1371/journal.pone.0102670
- 121 Savulescu J, Chalmers I, Blunt J. Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ* 1996;313:1390-3. doi:10.1136/bmj.313.7069.1390
- 122 Jia Y, Li B, Yang Z, et al. Trends of randomized clinical trials citing prior systematic reviews, 2007-2021. *JAMA Netw Open* 2023;6:e234219. doi:10.1001/jamanetworkopen.2023.4219
- 123 Bruce J, Mazuquin B, Canaway A, et al, Prevention of Shoulder Problems Trial (PROSPER) Study Group. Exercise versus usual care after non-reconstructive breast cancer surgery (UK PROSPER): multicentre randomised controlled trial and economic evaluation. *BMJ* 2021;375:e066542. doi:10.1136/bmj-2021-066542
- 124 AHRQ. Using the PICOTS Framework to Strengthen Evidence Gathered in Clinical Trials—Guidance from the AHRQ's Evidence-based Practice Centers Program. Available at: <https://www.fda.gov/media/109448/download> [accessed 13 March 2024].
- 125 Golder S, Loke YK, Wright K, Norman G. Reporting of adverse events in published and unpublished studies of health care interventions: a systematic review. *PLoS Med* 2016;13:e1002127. doi:10.1371/journal.pmed.1002127
- 126 Lineberry N, Berlin JA, Mansi B, et al. Recommendations to improve adverse event reporting in clinical trial publications: a joint pharmaceutical industry/journal editor perspective. *BMJ* 2016;355:i5078. doi:10.1136/bmj.i5078
- 127 Junqueira DR, Phillips R, Zorzelli L, et al. Time to improve the reporting of harms in randomized controlled trials. *J Clin Epidemiol* 2021;136:216-20. doi:10.1016/j.jclinepi.2021.04.020
- 128 Phillips R, Hazell L, Sauzet O, Cornelius V. Analysis and reporting of adverse events in randomised controlled trials: a review. *BMJ Open* 2019;9:e024537. doi:10.1136/bmjopen-2018-024537
- 129 Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012;308:2594-604. doi:10.1001/jama.2012.87802
- 130 Eldridge SM, Chan CL, Campbell MJ, et al, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ* 2016;355:i5239. doi:10.1136/bmj.i5239
- 131 Park JJH, Harari O, Dron L, Lester RT, Thorlund K, Mills EJ. An overview of platform trials with a checklist for clinical readers. *J Clin Epidemiol* 2020;125:1-8. doi:10.1016/j.jclinepi.2020.04.025
- 132 Schwartz D, Lelloch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Clin Epidemiol* 2009;62:499-505. doi:10.1016/j.jclinepi.2009.01.012
- 133 Fleming TR. Clinical trials: discerning hype from substance. *Ann Intern Med* 2010;153:400-6. doi:10.7326/0003-4819-153-6-201009210-00008
- 134 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
- 135 Kahan BC, Hindley J, Edwards M, Cro S, Morris TP. The estimands framework: a primer on the ICH E9(R1) addendum. *BMJ* 2024;384:e076316. doi:10.1136/bmj-2023-076316
- 136 European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Available at: [https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf). 2020.
- 137 Homer V, Yap C, Bond S, et al. Early phase clinical trials extension to guidelines for the content of statistical analysis plans. *BMJ* 2022;376:e068177. doi:10.1136/bmj-2021-068177
- 138 Kahan BC, Hall SS, Beller EM, et al. Reporting of factorial randomised trials: extension of the CONSORT 2010 statement. *JAMA* 2023;330:2106-14. doi:10.1001/jama.2023.19793
- 139 Kahan BC, Hall SS, Beller EM, et al. Consensus Statement for Protocols of Factorial Randomized Trials: Extension of the SPIRIT 2013 Statement. *JAMA Netw Open* 2023;6:e2346121. doi:10.1001/jamanetworkopen.2023.46121
- 140 Paskins Z, Bromley K, Lewis M, et al. Clinical effectiveness of one ultrasound guided intra-articular corticosteroid and local anaesthetic injection in addition to advice and education for hip osteoarthritis (HIT trial): single blind, parallel group, three arm, randomised controlled trial. *BMJ* 2022;377:e068446. doi:10.1136/bmj-2021-068446
- 141 Costa ML, Achten J, Ooms A, et al, DRAFFT2 Collaborators. Surgical fixation with K-wires versus casting in adults with fracture of distal radius: DRAFFT2 multicentre randomised clinical trial. *BMJ* 2022;376:e068041. doi:10.1136/bmj-2021-068041
- 142 Edwardson CL, Biddle SJH, Clemes SA, et al. Effectiveness of an intervention for reducing sitting time and improving health in office workers: three arm cluster randomised controlled trial. *BMJ* 2022;378:e069288. doi:10.1136/bmj-2021-069288

- 143 INSPIRATION-S Investigators. Atorvastatin versus placebo in patients with covid-19 in intensive care: randomized controlled trial. *BMJ* 2022;376:e068407.
- 144 Staley K. Exploring impact: public involvement in NHS, public health and social care research. INVOLVE, Eastleigh. <https://wellcomecollection.org/works/wtbgmgvr>. 2009.
- 145 Price A, Albarqouni L, Kirkpatrick J, et al. Patient and public involvement in the design of clinical trials: An overview of systematic reviews. *J Eval Clin Pract* 2018;24:240-53. doi:10.1111/jep.12805
- 146 National Institute for Health and Care Research. Briefing notes for researchers - public involvement in NHS, health and social care research (version 1.0) [Accessed 11 Sept 2023] 2021
- 147 Forsythe LP, Carman KL, Szydłowski V, et al. Patient engagement in research: early findings from the Patient-Centered Outcomes Research Institute. *Health Aff (Millwood)* 2019;38:359-67. doi:10.1377/hlthaff.2018.05067
- 148 Vanderhout S, Nicholls S, Monfaredi Z, et al. Facilitating and supporting the engagement of patients, families and caregivers in research: the "Ottawa model" for patient engagement in research. *Res Involv Engagem* 2022;8:25. doi:10.1186/s40900-022-00350-0
- 149 NHMRC. Guidelines for Guidelines: Consumer involvement. <https://www.nhmrc.gov.au/guidelinesforguidelines/plan/consumer-involvement>. 2018. [Accessed 11 Sept 2023]
- 150 Staniszewska S, Brett J, Mockford C, Barber R. The GRIPP checklist: strengthening the quality of patient and public involvement reporting in research. *Int J Technol Assess Health Care* 2011;27:391-9. doi:10.1017/S0266462311000481
- 151 Staniszewska S, Brett J, Simera I, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *BMJ* 2017;358:j3453. doi:10.1136/bmj.j3453
- 152 NHMRC. Statement on Consumer and Community involvement in Health and Medical Research. <https://www.nhmrc.gov.au/sites/default/files/documents/reports/consumer-community-involvement.pdf> [Accessed 09 Sept 2023]: Consumers Health Forum of Australia; 2016
- 153 CIHR-IRSC. Strategy for Patient-Oriented Research patient engagement framework. 2019. [https://cihr-irsc.gc.ca/e/documents/spor\\_framework-en.pdf](https://cihr-irsc.gc.ca/e/documents/spor_framework-en.pdf) [Accessed 09 Sept 2023]
- 154 Richards T, Godlee F. The BMJ's own patient journey. *BMJ* 2014;348:g3726. doi:10.1136/bmj.g3726
- 155 BMJ. New requirements for patient and public involvement statements in BMJ Open. 2018. <https://blogs.bmj.com/bmjopen/2018/03/23/new-requirements-for-patient-and-public-involvement-statements-in-bmj-open/> [Accessed 09 Sept 2023].
- 156 BMJ. BMJ author hub: patient and public partnership. <https://authors.bmj.com/policies/patient-public-partnership/> [Accessed 11 Sept 2023] [
- 157 Husson M, Dechartres A, Ramjee B, et al. Patient and public involvement is suboptimal in randomized controlled trials addressing a chronic condition. *J Clin Epidemiol* 2023;160:71-82. doi:10.1016/j.jclinepi.2023.06.012
- 158 de longh A, Severnwright A, Taylor J. Patient and public involvement statements in British journal of occupational therapy: An important step. *Br J Occup Ther* 2021;84:459-60. doi:10.1177/03080226211030103.
- 159 Research Involvement and Engagement: preparing your manuscript. <https://researchinvolvement.biomedcentral.com/submission-guidelines/preparing-your-manuscript/research> [Accessed 2023-10-19]
- 160 Davis S, Pinfold V, Catchpole J, Lovelock C, Senthil B, Kenny A. Reporting lived experience work. *Lancet Psychiatry* 2024;11:8-9. doi:10.1016/S2215-0366(23)00402-9
- 161 Gray R, Brasier C, Zirnsak T-M, Ng AH. Reporting of patient and public involvement and engagement (PPIE) in clinical trials published in nursing science journals: a descriptive study. *Res Involv Engagem* 2021;7:88. doi:10.1186/s40900-021-00331-9
- 162 Patel VA, Shelswell J, Hillyard N, Pavitt S, Barber SK. A study of the reporting of patient and public involvement and engagement (PPIE) in orthodontic research. *J Orthod* 2021;48:42-51. doi:10.1177/1465312520968574
- 163 Benizri N, Hallot S, Burns K, Goldfarb M. Patient and family representation in randomized clinical trials published in 3 medical and surgical journals: a systematic review. *JAMA Netw Open* 2022;5:e2230858. doi:10.1001/jamanetworkopen.2022.30858
- 164 Vanderhout S, Nevins P, Nicholls SG, et al. Patient and public involvement in pragmatic trials: online survey of corresponding authors of published trials. *CMAJ Open* 2023;11:E826-37. doi:10.9778/cmajo.202220198
- 165 Blumer JL, Findling RL, Shih WJ, Soubrane C, Reed MD. Controlled clinical trial of zolpidem for the treatment of insomnia associated with attention-deficit/ hyperactivity disorder in children 6 to 17 years of age. *Pediatrics* 2009;123:e770-6. doi:10.1542/peds.2008-2945
- 166 Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2019;321:1610-20. doi:10.1001/jama.2019.3087
- 167 Dwan K, Li T, Altman DG, Elbourne D. CONSORT 2010 statement: extension to randomised crossover trials. *BMJ* 2019;366:l4378. doi:10.1136/bmj.l4378
- 168 Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. CONSORT 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661. doi:10.1136/bmj.e5661
- 169 Odubayo A, Emdin CA, Hsiao AJ, et al. Association between trial registration and positive study findings: cross sectional study (Epidemiological Study of Randomized Trials-ESORT). *BMJ* 2017;356:j917. doi:10.1136/bmj.j917
- 170 Dargie HJ. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: the CAPRICORN randomised trial. *Lancet* 2001;357:1385-90. doi:10.1016/S0140-6736(00)04560-8
- 171 Santer M, Muller I, Becque T, et al. Eczema Care Online behavioural interventions to support self-care for children and young people: two independent, pragmatic, randomised controlled trials. *BMJ* 2022;379:e072007. doi:10.1136/bmj.2022-072007
- 172 Orkin AM, Gill PJ, Ghersi D, et al. CONSERVE Group. Guidelines for Reporting Trial Protocols and Completed Trials Modified Due to the COVID-19 Pandemic and Other Extenuating Circumstances: The CONSERVE 2021 Statement. *JAMA* 2021;326:257-65. doi:10.1001/jama.2021.9941
- 173 Grant AM, Altman DG, Babiker AB, et al. DAMOCLES study group. Issues in data monitoring and interim analysis of trials. *Health Technol Assess* 2005;9:i-238, iii-iv. doi:10.3310/hta9070
- 174 Dimairo M, Pallmann P, Wason J, et al. ACE Consensus Group. The Adaptive designs CONSORT Extension (ACE) statement: a checklist with explanation and elaboration guideline for reporting randomised trials that use an adaptive design. *BMJ* 2020;369:m115. doi:10.1136/bmj.m115
- 175 Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004;171:735-40. doi:10.1503/cmaj.104.1086
- 176 TARG Meta-Research Group & Collaborators. Estimating the prevalence of discrepancies between study registrations and publications: a systematic review and meta-analyses. *BMJ Open* 2023;13:e076264. doi:10.1136/bmjopen-2023-076264
- 177 Blair JC, McKay A, Ridyard C, et al. SCIPi Investigators. Continuous subcutaneous insulin infusion versus multiple daily injection regimens in children and young people at diagnosis of type 1 diabetes: pragmatic randomised controlled trial and economic evaluation. *BMJ* 2019;365:l1226. doi:10.1136/bmj.l1226
- 178 Tappin D, Bauld L, Purves D, et al. Cessation in Pregnancy Incentives Trial Team. Financial incentives for smoking cessation in pregnancy: randomised controlled trial. *BMJ* 2015;350:h134. doi:10.1136/bmj.h134
- 179 Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply"? *Lancet* 2005;365:82-93. doi:10.1016/S0140-6736(04)17670-8
- 180 Weiss NS, Koepsell TD, Psaty BM. Generalizability of the results of randomized trials. *Arch Intern Med* 2008;168:133-5. doi:10.1001/archinternmed.2007.30
- 181 Hopewell S, Keene DJ, Marian IR, et al. GRASP Trial Group. Progressive exercise compared with best practice advice, with or without corticosteroid injection, for the treatment of patients with rotator cuff disorders (GRASP): a multicentre, pragmatic, 2 x 2 factorial, randomised controlled trial. *Lancet* 2021;398:416-28. doi:10.1016/S0140-6736(21)00846-1
- 182 Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. *Res Integr Peer Rev* 2016;1:2. doi:10.1186/s41073-016-0007-6
- 183 Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA* 2007;297:1233-40. doi:10.1001/jama.297.11.1233
- 184 Blümle A, Meerpolhl J, Rücker G, Antes G, Schumacher M, von Elm E. Reporting of eligibility criteria of randomised trials: cohort study comparing trial protocols with subsequent articles. *BMJ* 2011;342:d1828. doi:10.1136/bmj.d1828
- 185 Shapiro SH, Weijer C, Freedman B. Reporting the study populations of clinical trials. Clear transmission or static on the line? *Clin Epidemiol* 2000;53:973-9. doi:10.1016/S0895-4356(00)00227-4
- 186 Gandhi M, Ameli N, Bacchetti P, et al. Eligibility criteria for HIV clinical trials and generalizability of results: the gap between published reports and study protocols. *AIDS* 2005;19:1885-96. doi:10.1097/01.aids.0000189866.67182.f7

## RESEARCH METHODS AND REPORTING

- 187 Molyneux A, Kerr R, Stratton I, et al. International Subarachnoid Aneurysm Trial (ISAT) Collaborative Group. International Subarachnoid Aneurysm Trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised trial. *Lancet* 2002;360:1267-74. doi:10.1016/S0140-6736(02)11314-6
- 188 Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, CONSORT Group. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008;148:295-309. doi:10.7326/0003-4819-148-4-200802190-00008
- 189 Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Intern Med* 2002;137:511-20. doi:10.7326/0003-4819-137-6-200209170-00012
- 190 Devereaux PJ, Bhандari M, Clarke M, et al. Need for expertise based randomised controlled trials. *BMJ* 2005;330:88. doi:10.1136/bmj.330.7482.88
- 191 McGrath PD, Wennberg DE, Dickens JD Jr. Relation between operator and hospital volume and outcomes following percutaneous coronary interventions in the era of the coronary stent. *JAMA* 2000;284:3139-44. doi:10.1001/jama.284.24.3139
- 192 Hodgson DC, Zhang W, Zaslavsky AM, Fuchs CS, Wright WE, Ayanian JZ. Relation of hospital volume to colostomy rates and survival for patients with rectal cancer. *J Natl Cancer Inst* 2003;95:708-16. doi:10.1093/jnci/95.10.708
- 193 Khuri SF, Daley J, Henderson W, et al. Relation of surgical volume to outcome in eight common operations: results from the VA National Surgical Quality Improvement Program. *Ann Surg* 1999;230:414-29, discussion 429-32. doi:10.1097/00000658-199909000-00014
- 194 Urbach DR, Baxter NN. Does it matter what a hospital is "high volume" for? Specificity of hospital volume-outcome associations for surgical procedures: analysis of administrative data. *Qual Saf Health Care* 2004;13:379-83. doi:10.1136/qhc.13.5.379
- 195 Jacquier I, Boutron I, Moher D, Roy C, Ravaud P. The reporting of randomized clinical trials using a surgical intervention is in need of immediate improvement: a systematic review. *Ann Surg* 2006;244:677-83. doi:10.1097/01.sla.0000242707.44007.80
- 196 Madanitsa M, Barsosio HC, Minja DTR, et al. Effect of monthly intermittent preventive treatment with dihydroartemisinin-piperaquine with and without azithromycin versus monthly sulfadoxine-pyrimethamine on adverse pregnancy outcomes in Africa: a double-blind randomised, partly placebo-controlled trial. *Lancet* 2023;401:1020-36. doi:10.1016/S0140-6736(22)02535-1
- 197 Nguyen C, Boutron I, Zegarra-Parodi R, et al. Effect of osteopathic manipulative treatment vs sham treatment on activity limitations in patients with nonspecific subacute and chronic low back pain: a randomized clinical trial. *JAMA Intern Med* 2021;181:620-30. doi:10.1001/jamainternmed.2021.0005
- 198 Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008;336:1472-4. doi:10.1136/bmjj.39590.732037.47
- 199 Duff JM, Leather H, Walden EO, LaPlant KD, George TJ Jr. Adequacy of published oncology randomized controlled trials to provide therapeutic details needed for clinical application. *J Natl Cancer Inst* 2010;102:702-5. doi:10.1093/jnci/djq117
- 200 Schroter S, Glasziou P, Heneghan C. Quality of descriptions of treatments: a review of published randomised controlled trials. *BMJ Open* 2012;2:e001978. doi:10.1136/bmjjopen-2012-001978
- 201 Hoffmann TC, Erueti C, Glasziou PP. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. *BMJ* 2013;347:f3755. doi:10.1136/bmjj.f3755
- 202 Abell B, Glasziou P, Hoffmann T. Reporting and replicating trials of exercise-based cardiac rehabilitation: do we know what the researchers actually did? *Circ Cardiovasc Qual Outcomes* 2015;8:187-94. doi:10.1161/CIRCOUTCOMES.114.001381
- 203 Ndounga Diakou LA, Ntoumi F, Ravaud P, Boutron I. Avoidable waste related to inadequate methods and incomplete reporting of interventions: a systematic review of randomized trials performed in Sub-Saharan Africa. *Trials* 2017;18:291. doi:10.1186/s13063-017-2034-0
- 204 Golomb BA, Erickson LC, Koperski S, Sack D, Enkin M, Howick J. What's in placebos: who knows? Analysis of randomized, controlled trials. *Ann Intern Med* 2010;153:532-5. doi:10.7326/0003-4819-153-8-201010190-00010
- 205 Yu AM, Balasubramaniam B, Offringa M, Kelly LE. Reporting of interventions and "standard of care" control arms in pediatric clinical trials: a quantitative analysis. *Pediatr Res* 2018;84:393-8. doi:10.1038/s41390-018-0019-7
- 206 Sanders S, Gibson E, Glasziou P, Hoffmann T. Nondrug interventions for reducing SARS-CoV-2 transmission are frequently incompletely reported. *J Clin Epidemiol* 2023;157:102-9. doi:10.1016/j.jclinepi.2023.02.006
- 207 Howick J, Webster RK, Rees JL, et al. TiDieR-Placebo: A guide and checklist for reporting placebo and sham controls. *PLoS Med* 2020;17:e1003294. doi:10.1371/journal.pmed.1003294
- 208 Phillips WR, Sturgiss E, Hunik L, et al. Improving the Reporting of Primary Care Research: An International Survey of Researchers. *J Am Board Fam Med* 2021;34:12-21. doi:10.3122/jabfm.2021.01.200266
- 209 Dodd S, White IR, Williamson P. Nonadherence to treatment protocol in published randomised controlled trials: a review. *Trials* 2012;13:84. doi:10.1186/1745-6215-13-84
- 210 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898. doi:10.1136/bmj.l4898
- 211 Higgins J, Sterne J, Savović J, et al. A revised tool for assessing risk of bias in randomized trials. In: Chandler J, McKenzie J, Boutron I, Welch V, eds. *Cochrane Methods Cochrane Database of Systematic Reviews 2016, Issue 10 (Suppl 1)*. 2016: 29-32.
- 212 Zuidgeest MGP, Welsing PMJ, van Thiel GJMW, et al. WP3 of the GetReal consortium. Series: Pragmatic trials and real world evidence: Paper 5. Usual care and real life comparators. *J Clin Epidemiol* 2017;90:92-8. doi:10.1016/j.jclinepi.2017.07.001
- 213 Turner KM, Huntley A, Yardley T, Dawson S, Dawson S. Defining usual care comparators when designing pragmatic trials of complex health interventions: a methodology review. *Trials* 2024;25:117. doi:10.1186/s13063-024-07956-7
- 214 Webster RK, Howick J, Hoffmann T, et al. Inadequate description of placebo and sham controls in a systematic review of recent trials. *Eur J Clin Invest* 2019;49:e13169. doi:10.1111/eci.13169
- 215 Paterson C, Dieppe P. Characteristic and incidental (placebo) effects in complex interventions such as acupuncture. *BMJ* 2005;330:1202-5. doi:10.1136/bmj.330.7501.1202
- 216 Bannuru RR, McAlindon TE, Sullivan MC, Wong JB, Kent DM, Schmid CH. Effectiveness and Implications of Alternative Placebo Treatments: A Systematic Review and Network Meta-analysis of Osteoarthritis Trials. *Ann Intern Med* 2015;163:365-72. doi:10.7326/M15-0623
- 217 COMET. Core Outcome Measures in Effectiveness Trials. <https://comet-initiative.org/> [Accessed 6 Nov 2023]
- 218 Hughes KL, Clarke M, Williamson PR. A systematic review finds Core Outcome Set uptake varies widely across different areas of health. *J Clin Epidemiol* 2021;129:114-23. doi:10.1016/j.jclinepi.2020.09.029
- 219 Zarin DA, Tse T, Williams RJ, Calif RM, Ide NC. The ClinicalTrials.gov results database--update and key issues. *N Engl J Med* 2011;364:852-60. doi:10.1056/NEJMsa1012065
- 220 Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005;365:1159-62. doi:10.1016/S0140-6736(05)71879-1
- 221 Hopewell S, Dutton S, Yu L-M, Chan A-W, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ* 2010;340:c723. doi:10.1136/bmj.c723
- 222 Bridgman AC, McPhie ML, Voineskos SH, Chan A-W, Drucker AM. Reporting of primary outcome measures and sample size calculations in randomized controlled trials in dermatology journals. *J Am Acad Dermatol* 2022;87:912-4. doi:10.1016/j.jaad.2021.12.022
- 223 Stoll M, Lindner S, Marquardt B, et al. Completeness and consistency of primary outcome reporting in COVID-19 publications in the early pandemic phase: a descriptive study. *BMC Med Res Methodol* 2023;23:173. doi:10.1186/s12874-023-01991-9
- 224 McDowell I. *Measuring health: a guide to rating scales and questionnaires*. 3rd ed. Oxford University Press, 2006. doi:10.1093/acprof:oso/9780195165678.001.0001.
- 225 Streiner DL, Norman GR, Cairney J. *Health measurement scales: a practical guide to their development and use*. Oxford University Press; 2015.
- 226 Sanders C, Egger M, Donovan J, Tallon D, Frankel S. Reporting on quality of life in randomised controlled trials: bibliographic study. *BMJ* 1998;317:1191-4. doi:10.1136/bmjj.317.7167.1191
- 227 Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007;8:39. doi:10.1186/1745-6215-8-39
- 228 Mc Cord KA, Ewald H, Agarwal A, et al. Treatment effects in randomised trials using routinely collected data for outcome assessment versus traditional trials: meta-research study. *BMJ* 2021;372:n450. doi:10.1136/bmjj.n450
- 229 Torre-Cisneros J, San-Juan R, Rosso-Fernández CM, et al. Tuberculosis prophylaxis with levofloxacin in liver transplant patients is associated with a high incidence of tenosynovitis: safety analysis of a multicenter randomized trial. *Clin Infect Dis* 2015;60:1642-9. doi:10.1093/cid/civ156
- 230 Mohraz M, Vahdat K, Ghamari SH, et al. Efficacy and safety of an inactivated virus-particle vaccine for SARS-CoV-2, BIV1-Coviran: randomised, placebo controlled, double blind, multicentre, phase 3 clinical trial. *BMJ* 2023;382:e070464. doi:10.1136/bmjj-2023-070464

- 231 Qureshi R, Mayo-Wilson E, Li T. Harms in Systematic Reviews Paper 1: An introduction to research on harms. *J Clin Epidemiol* 2022;143:186-96. doi:10.1016/j.jclinepi.2021.10.023
- 232 Manson JE, Chlebowski RT, Stefanick ML, et al. Menopausal hormone therapy and health outcomes during the intervention and extended poststopping phases of the Women's Health Initiative randomized trials. *JAMA* 2013;310:1353-68. doi:10.1001/jama.2013.278040
- 233 Mayo-Wilson E, Fusco N, Li T, Hong H, Canner JK, Dickersin K, MUDS investigators. Harms are assessed inconsistently and reported inadequately Part 2: nonsystematic adverse events. *J Clin Epidemiol* 2019;113:11-9. doi:10.1016/j.jclinepi.2019.04.020
- 234 Bagul NB, Kirkham JJ. The reporting of harms in randomized controlled trials of hypertension using the CONSORT criteria for harm reporting. *Clin Exp Hypertens* 2012;34:548-54. doi:10.3109/10641963.2012.681724
- 235 Qureshi R, Mayo-Wilson E, Rittiphairoj T, McAdams-DeMarco M, Guallar E, Li T. Harms in Systematic Reviews Paper 2: Methods used to assess harms are neglected in systematic reviews of gabapentin. *J Clin Epidemiol* 2022;143:212-23. doi:10.1016/j.jclinepi.2021.10.024
- 236 Qureshi R, Mayo-Wilson E, Rittiphairoj T, McAdams-DeMarco M, Guallar E, Li T. Harms in Systematic Reviews Paper 3: Given the same data sources, systematic reviews of gabapentin have different results for harms. *J Clin Epidemiol* 2022;143:224-41. doi:10.1016/j.jclinepi.2021.10.025
- 237 Stephens MDB, Routledge P, Talbot JCC. *Detection of new adverse drug reactions*. 4th ed. Macmillan Basingstoke, 1998.
- 238 Levit LA, Perez RP, Smith DC, Schilsky RL, Hayes DF, Vose JM. Streamlining adverse events reporting in oncology: an american society of clinical oncology research statement. *J Clin Oncol* 2018;36:617-23. doi:10.1200/JCO.2017.75.8193
- 239 Taieb J, Bouche O, André T, et al, SAMCO-PRODIGE 54 Investigators. Avelumab vs standard second-line chemotherapy in patients with metastatic colorectal cancer and microsatellite instability: a randomized clinical trial. *JAMA Oncol* 2023;9:1356-63. doi:10.1001/jamaoncol.2023.2761
- 240 Casswell Ej, Cro S, Cornelius VR, et al, ASCOT Investigator Study Group. Randomised controlled trial of adjunctive triamcinolone acetonide in eyes undergoing vitreoretinal surgery following open globe trauma: The ASCOT study. *Br J Ophthalmol* 2024;108:440-8. doi:10.1136/bjo-2022-322787
- 241 Khunti K, Highton PJ, Waheed G, et al. Promoting physical activity with self-management support for those with multimorbidity: a randomised controlled trial. *Br J Gen Pract* 2021;71:e921-30. doi:10.3399/BJGP.2021.0172
- 242 Palmer AJR, Ayyar Gupta V, Fernquest S, et al, FAIT Study Group. Arthroscopic hip surgery compared with physiotherapy and activity modification for the treatment of symptomatic femoroacetabular impingement: multicentre randomised controlled trial. *BMJ* 2019;364:i185. doi:10.1136/bmj.i185
- 243 Campbell MJ, Julius SA, Altman DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *BMJ* 1995;311:1145-8. doi:10.1136/bmj.311.7013.1145
- 244 Cook JA, Julius SA, Sones W, et al. DELTA<sup>2</sup> guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. *BMJ* 2018;363:k3750. doi:10.1136/bmj.k3750
- 245 Pandis N, Chung B, Scherer RW, Elbourne D, Altman DG. CONSORT 2010 statement: extension checklist for reporting within person randomised trials. *BMJ* 2017;357:j2835. doi:10.1136/bmj.j2835
- 246 Speich B. Adequate reporting of the sample size calculation in surgical randomized controlled trials. *Surgery* 2020;167:812-4. doi:10.1016/j.surg.2019.10.011
- 247 Weinberg T, Wang G, Lam K, Kitchen J, Chan A-W. Reporting of sample-size calculations for randomized trials in dermatology: comparison of publications with registries. *Br J Dermatol* 2019;180:929-30. doi:10.1111/bjd.17332
- 248 Amiri M, Kumbhare D. Randomized controlled trials in non-pharmacological rehabilitation research: a scoping review of the reporting of sample size calculation, randomization procedure, and statistical analyses. *Eur J Phys Rehabil Med* 2020;56:790-8.
- 249 Dumbrigue HB, Dumbrigue EC, Dumbrigue DC, Chingbingyong MI. Reporting of sample size parameters in randomized controlled trials published in prosthodontic journals. *J Prosthodont* 2019;28:159-62. doi:10.1111/jopr.13010
- 250 Nontshe M, Khan S, Mandebvu T, Merrifield B, Rodseth RN. Sample-size determination and adherence in randomised controlled trials published in anaesthesia journals. *Southern Afr J Anaesth Analg* 2018;24:40-6. doi:10.1080/220201181.2018.1439602 .
- 251 Copsey B, Thompson JY, Vadher K, et al. Sample size calculations are poorly conducted and reported in many randomized trials of hip and knee osteoarthritis: results of a systematic review. *J Clin Epidemiol* 2018;104:52-61. doi:10.1016/j.jclinepi.2018.08.013
- 252 Lee PH, Tse ACY. The quality of the reported sample size calculations in randomized controlled trials indexed in PubMed. *Eur J Intern Med* 2017;40:16-21. doi:10.1016/j.ejim.2016.10.008
- 253 Althouse AD. Post hoc power: not empowering, just misleading. *J Surg Res* 2021;259:A3-6. doi:10.1016/j.jss.2019.10.049
- 254 Goodman SN, Berlin JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Ann Intern Med* 1994;121:200-6. doi:10.7326/0003-4819-121-3-199408010-00008
- 255 Horne AW, Tong S, Moakes CA, et al, GEM3 collaborative. Combination of gefitinib and methotrexate to treat tubal ectopic pregnancy (GEM3): a multicentre, randomised, double-blind, placebo-controlled trial. *Lancet* 2023;401:655-63. doi:10.1016/S0140-6736(22)02478-3
- 256 Hall DA, Ray J, Watson J, et al. A balanced randomised placebo controlled blinded phase IIa multi-centre study to investigate the efficacy and safety of AUTO0063 versus placebo in subjective tinnitus: The QUIET-1 trial. *Hear Res* 2019;377:153-66. doi:10.1016/j.heares.2019.03.018
- 257 Hogan LE, Brown PA, Ji L, et al. Children's Oncology Group AALL1331: phase III trial of blinatumomab in children, adolescents, and young adults with low-risk B-cell ALL in first relapse. *J Clin Oncol* 2023;41:4118-29. doi:10.1200/JCO.22.02200
- 258 Reignier J, Plantefève G, Mira JP, et al, NUTRIREA-3 Trial InvestigatorsClinical Research in Intensive CareSepsis (CRICS-TRIGGERSEP)Group. Low versus standard calorie and protein feeding in ventilated adults with shock: a randomised, controlled, multicentre, open-label, parallel-group trial (NUTRIREA-3). *Lancet Respir Med* 2023;11:602-12. doi:10.1016/S2213-2600(23)00092-9
- 259 Ciolino JD, Kaizer AM, Bonner LB. Guidance on interim analysis methods in clinical trials. *J Clin Transl Sci* 2023;7:e124. doi:10.1017/cts.2023.552
- 260 Stegert M, Kasenda B, von Elm E, et al, DISCO study group. An analysis of protocols and publications suggested that most discontinuations of clinical trials were not based on preplanned interim analyses or stopping rules. *J Clin Epidemiol* 2016;69:152-60. doi:10.1016/j.jclinepi.2015.05.023
- 261 Padoei F, Mamsharifi P, Hazeghe P, et al. The therapeutic effect of N-acetylcysteine as an add-on to methadone maintenance therapy medication in outpatients with substance use disorders: A randomized, double-blind, placebo-controlled clinical trial. *Brain Behav* 2023;13:e2823. doi:10.1002/brb3.2823
- 262 Karyotaki E, Klein AM, Ciharova M, et al. Guided internet-based transdiagnostic individually tailored Cognitive Behavioral Therapy for symptoms of depression and/or anxiety in college students: A randomized controlled trial. *Behav Res Ther* 2022;150:104028. doi:10.1016/j.brat.2021.104028
- 263 Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. *BMJ* 1948;2:769-82. doi:10.1136/bmj.2.4582.769
- 264 Schulz KF. Randomized controlled trials. *Clin Obstet Gynecol* 1998;41:245-56. doi:10.1097/00003081-199806000-00005
- 265 Greenland S. Randomization, statistics, and causal inference. *Epidemiology* 1990;1:421-9. doi:10.1097/00001648-199011000-00003
- 266 Armitage P. The role of randomization in clinical trials. *Stat Med* 1982;1:349-52. doi:10.1002/sim.4780010412
- 267 Kleijnen J, Götzsche P, Kunz RA, Oxman A, Chalmers I. So what's so special about randomisation? In: Manyard A, Chalmers I, eds. *Non-random reflections on health services research*. BMJ Publishing Group, 1997: 93-106.
- 268 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12. doi:10.1001/jama.1995.03520290060030
- 269 Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994;272:125-8. doi:10.1001/jama.1994.03520020051014
- 270 Chalmers I. Assembling comparison groups to assess the effects of health care. *J R Soc Med* 1997;90:379-86. doi:10.1177/014107689709000706
- 271 Altman DG. Randomisation. *BMJ* 1991;302:1481-2. doi:10.1136/bmj.302.6791.1481
- 272 Moher D. CONSORT: an evolving tool to help improve the quality of reports of randomized controlled trials. Consolidated Standards of Reporting Trials. *JAMA* 1998;279:1489-91. doi:10.1001/jama.279.18.1489
- 273 Kjaergard L, Villumsen J, Gluud C. Quality of randomised clinical trials affects estimates of intervention efficacy. The best evidence for health care: the role of The Cochrane Collaboration - abstracts of the 7th Cochrane Colloquium; 1999 5-9 Oct; Rome, Italy 1999.

- 274 Juni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic Reviews in Health Care: Meta-analysis in Context*. BMJ Books, 2001. doi:10.1002/9780470693926.ch5.
- 275 Savović J, Jones H, Altman D, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technol Assess* 2012;16:1-82. doi:10.3310/hta16350
- 276 Savović J, Turner RM, Mawdsley D, et al. Association between risk-of-bias assessments and results of randomized trials in Cochrane reviews: the ROBES meta-epidemiologic study. *Am J Epidemiol* 2018;187:1113-22. doi:10.1093/aje/kwx344
- 277 Saltaji H, Armijo-Olivo S, Cummings GG, Amin M, da Costa BR, Flores-Mir C. Impact of selection bias on treatment effect size estimates in randomized trials of oral health interventions: a meta-epidemiological study. *J Dent Res* 2018;97:5-13. doi:10.1177/0022034517725049
- 278 Armijo-Olivo S, Saltaji H, da Costa BR, Fuentes J, Ha C, Cummings GG. What is the influence of randomisation sequence generation and allocation concealment on treatment effects of physical therapy trials? A meta-epidemiological study. *BMJ Open* 2015;5:e008562. doi:10.1136/bmjjopen-2015-008562
- 279 Barcot O, Boric M, Poklepovic Pericic T, et al. Risk of bias judgments for random sequence generation in Cochrane systematic reviews were frequently not in line with Cochrane Handbook. *BMC Med Res Methodol* 2019;19:170. doi:10.1186/s12874-019-0804-y
- 280 Mentz RJ, Anstrom KJ, Eisenstein EL, et al. TRANSFORM-HF Investigators. Effect of torsemide vs furosemide after discharge on all-cause mortality in patients hospitalized with heart failure: the TRANSFORM-HF randomized clinical trial. *JAMA* 2023;329:214-23. doi:10.1001/jama.2022.23924
- 281 Statema-Lohmeijer CH, Schats R, Lissenberg-Witte BI, Kosteljik EH, Lambalk CB, Vergouw CG. A short versus a long time interval between semen collection and intrauterine insemination: a randomized controlled clinical trial. *Hum Reprod* 2023;38:811-9. doi:10.1093/humrep/dead044
- 282 Reichenbach S, Jüni P, Hincapíe CA, et al. Effect of transcutaneous electrical nerve stimulation (TENS) on knee pain and physical function in patients with symptomatic knee osteoarthritis: the ETRELKA randomized clinical trial. *Osteoarthritis Cartilage* 2022;30:426-35. doi:10.1016/j.joca.2021.10.015
- 283 Creinin MD, Meyn LA, Borgatta L, et al. Multicenter comparison of the contraceptive ring and patch: a randomized controlled trial. *Obstet Gynecol* 2008;111:267-77. doi:10.1097/AOG.00000298338.58511.d1
- 284 Lachin JM. Properties of simple randomization in clinical trials. *Control Clin Trials* 1988;9:312-26. doi:10.1016/0197-2456(88)90046-3
- 285 Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer* 1976;34:585-612. doi:10.1038/bjc.1976.220
- 286 Schulz K, Grimes DA. *Essential concepts in clinical research: randomized controlled trials and observational epidemiology*. 2nd ed. Elsevier, 2019.
- 287 Altman DG, Bland JM. Statistics notes: how to randomise. *BMJ* 1999;319:703-4. doi:10.1136/bmj.319.7211.703
- 288 Schulz KF. Subverting randomization in controlled trials. *JAMA* 1995;274:1456-8. doi:10.1001/jama.1995.03530180050029
- 289 Enas GG, Enas NH, Spradlin CT, Wilson MG, Wiltse CG. Baseline comparability in clinical trials: prevention of "poststudy anxiety". *Drug Inf J* 1990;24:541-8. doi:10.1177/009286159002400312.
- 290 Pocock SJ. *Clinical trials: a practical approach*. John Wiley & Sons Ltd, 1983.
- 291 Treasure T, MacRae KD. Minimisation: the platinum standard for trials? Randomisation doesn't guarantee similarity of groups; minimisation does. *BMJ* 1998;317:362-3. doi:10.1136/bmj.317.7155.362
- 292 Khan BC, Rehal S, Cro S. Risk of selection bias in randomised trials. *Trials* 2015;16:405. doi:10.1186/s13063-015-0920-x
- 293 Clark L, Burke L, Margaret Carr R, Coleman E, Roberts G, Torgerson DJ. A review found small variable blocking schemes may not protect against selection bias in randomized controlled trials. *J Clin Epidemiol* 2022;141:90-8. doi:10.1016/j.jclinepi.2021.09.009
- 294 Altman DG, Doré CJ. Randomisation and baseline comparisons in clinical trials. *Lancet* 1990;335:149-53. doi:10.1016/0140-6736(90)90014-V
- 295 Adetugbo K, Williams H. How well are randomized controlled trials reported in the dermatology literature? *Arch Dermatol* 2000;136:381-5. doi:10.1001/archderm.136.3.381
- 296 Ciolino JD, Palac HL, Yang A, Vacca M, Belli HM. Ideal vs. real: a systematic review on handling covariates in randomized controlled trials. *BMC Med Res Methodol* 2019;19:136. doi:10.1186/s12874-019-0787-8
- 297 Sandborn WJ, Vermeire S, Peyrin-Biroulet L, et al. Etrasimod as induction and maintenance therapy for ulcerative colitis (ELEVATE): two randomised, double-blind, placebo-controlled, phase 3 studies. *Lancet* 2023;401:1159-71. doi:10.1016/S0140-6736(23)00061-2
- 298 Rai AV, Naik BD. Evaluation of remineralizing effect of zinc-carbonate hydroxyapatite on the reduction of postrestorative sensitivity: A randomized controlled clinical trial. *J Conserv Dent* 2023;26:56-66. doi:10.4103/jcd.jcd\_300\_22
- 299 Mendon A FÍT, Barreto Filho JH, Hungria MBCS, Magalhães TC. Efficacy of a single dose of esmolol to prevent extubation-related complications during emergence from anaesthesia: a randomized, double-blind, placebo-controlled trial. *Braz J Anesthesiol* 2023;73:426-33.
- 300 Rattanathanya T, Adulkasem N, Wongcharoenwata J, Ariyawatkul T, Chotigavanichaya C, Eamsobhana P. Perioperative blood loss reduction using a sterile exsanguination tourniquet for orthopedic femoral-related surgeries in children: a randomized controlled study. *J Orthop Surg Res* 2023;18:580. doi:10.1186/s13018-023-04046-3
- 301 Radford JA, Landorf KB, Buchbinder R, Cook C. Effectiveness of low-Dye taping for the short-term treatment of plantar heel pain: a randomised trial. *BMC Musculoskelet Disord* 2006;7:64. doi:10.1186/1471-2474-7-64
- 302 Chalmers TC, Levin H, Sacks HS, Reitman D, Berrier J, Nagalingam R. Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials. *Stat Med* 1987;6:315-28. doi:10.1002/sim.4780060320
- 303 Pocock SJ. Statistical aspects of clinical trial design. *Statistician* 2018;31:1-18. doi:10.2307/2988097.
- 304 Piaggio G, Elbourne D, Schulz KF, Villar J, Pinol AP, Gürmezoglu AM, WHO Research Group to evaluate Misoprostol in the Management of the Third Stage of Labour. The reporting of methods for reducing and detecting bias: an example from the WHO Misoprostol Third Stage of Labour equivalence randomised controlled trial. *BMC Med Res Methodol* 2003;3:19. doi:10.1186/1471-2288-3-19
- 305 Clark L, Mitchell N, Hewitt C, Torgerson D. Exploring engagement with authors of randomised controlled trials to develop recommendations to improve allocation concealment implementation and reporting. *F1000 Res* 2021;10:83. doi:10.12688/f1000research.27918.1.
- 306 Clark L, Schmidt U, Tharmanathan P, Adamson J, Hewitt C, Torgerson D. Poor reporting quality of key Randomization and Allocation Concealment details is still prevalent among published RCTs in 2011: a review. *J Eval Clin Pract* 2013;19:703-7. doi:10.1111/jep.12031
- 307 Pildal J, Hróbjartsson A, Jørgensen KJ, Hilden J, Altman DG, Götzsche PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol* 2007;36:847-57. doi:10.1093/ije/dym087
- 308 Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601-5. doi:10.1136/bmj.39465.451748.AD
- 309 Wang Y, Parpia S, Couban R, et al. Compelling evidence from meta-epidemiological studies demonstrates overestimation of effects in randomized trials that fail to optimize randomization and blind patients and outcome assessors. *J Clin Epidemiol* 2024;165:111211. PubMeddoi:10.1016/j.jclinepi.2023.11.001
- 310 Götzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989;10:31-56. doi:10.1016/0197-2456(89)90017-2
- 311 Mhashkar R, Djulbegovic B, Magazin A, Soares HP, Kumar A. Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. *J Clin Epidemiol* 2012;65:602-9. doi:10.1016/j.jclinepi.2011.10.016
- 312 Clark L, Schmidt U, Tharmanathan P, Adamson J, Hewitt C, Torgerson D. Allocation concealment: a methodological review. *J Eval Clin Pract* 2013;19:708-12. doi:10.1111/jep.12032
- 313 He J, Du L, Liu G, et al. Quality assessment of reporting of randomization, allocation concealment, and blinding in traditional Chinese medicine RCTs: a review of 3159 RCTs identified from 260 systematic reviews. *Trials* 2011;12:122. doi:10.1186/1745-6215-12-122
- 314 Zheng SL, Chan FT, Maclean E, Jayakumar S, Nabeboccus AA. Reporting trends of randomised controlled trials in heart failure with preserved ejection fraction: a systematic review. *Open Heart* 2016;3:e000449. doi:10.1136/openhrt-2016-000449
- 315 Kilicoglu H, Jiang L, Hoang L, Mayo-Wilson E, Vinkers CH, Otte WM. Methodology reporting improved over time in 176,469 randomized controlled trials. *J Clin Epidemiol* 2023;162:19-28. doi:10.1016/j.jclinepi.2023.08.004
- 316 Binks MJ, Bleakley AS, Pizzutto SJ, et al. Randomised controlled trial of perinatal vitamin D supplementation to prevent early-onset acute respiratory infections among Australian First Nations children: the 'D-Kids' study protocol. *BMJ Open Respir Res* 2023;10:e001646. doi:10.1136/bmjjresp-2023-001646

- 317 Jones CMP, Day RO, Koes BW, et al. OPAL Investigators Coordinators. Opioid analgesia for acute low back pain and neck pain (the OPAL trial): a randomised placebo-controlled trial. *Lancet* 2023;402:304-12. doi:10.1016/S0140-6736(23)00404-X
- 318 Evans JT, Stanesby O, Blizzard L, et al. trips4health: a single-blinded randomised controlled trial incentivising adult public transport use for physical activity gain. *Int J Behav Nutr Phys Act* 2023;20:98. doi:10.1186/s12966-023-01500-7
- 319 Webster J, Clarke S, Paterson D, et al. Routine care of peripheral intravenous catheters versus clinically indicated replacement: randomised controlled trial. *BMJ* 2008;337:a339. doi:10.1136/bmj.a339
- 320 Smith SA, Shah ND, Bryant SC, et al. Evidens Research Group. Chronic care model and shared care in diabetes: randomized trial of an electronic decision support system. *Mayo Clin Proc* 2008;83:747-57. doi:10.4065/83.7.747
- 321 Sacks FM, Bray GA, Carey VJ, et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N Engl J Med* 2009;360:859-73. doi:10.1056/NEJMoa0804748
- 322 Kapchuk TJ. Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bull Hist Med* 1998;72:389-433. doi:10.1353/bhm.1998.0159
- 323 Devereaux PJ, Manns BJ, Ghali WA, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001;285:2000-3. doi:10.1001/jama.285.15.2000
- 324 Haahr MT, Hróbjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. *Clin Trials* 2006;3:360-5. doi:10.1177/1740774506069153
- 325 Götzsche PC. Blinding during data analysis and writing of manuscripts. *Control Clin Trials* 1996;17:285-90, discussion 290-3. doi:10.1016/0197-2456(95)00263-4
- 326 Meinert CL. Masked monitoring in clinical trials--blind stupidity? *N Engl J Med* 1998;338:1381-2. doi:10.1056/NEJM199805073381911
- 327 Guyatt GH, Pugsley SO, Sullivan MJ, et al. Effect of encouragement on walking test performance. *Thorax* 1984;39:818-22. doi:10.1136/thx.39.11.818
- 328 Karolinski TR, Chalmers TC, Frenkel LD, Kapikian AZ, Lewis TL, Lynch JM. Ascorbic acid for the common cold. A prophylactic and therapeutic trial. *JAMA* 1975;231:1038-42. doi:10.1001/jama.1975.03240220018013
- 329 Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994;44:16-20. doi:10.1212/WNL.44.1.16
- 330 Janiaud P, Dal-Ré R, Ioannidis JPA. Assessment of Pragmatism in Recently Published Randomized Clinical Trials. *JAMA Intern Med* 2018;178:1278-80. doi:10.1001/jamainternmed.2018.3321
- 331 Carley SD, Libetta C, Flavin B, Butler J, Tong N, Sammy I. An open prospective randomised trial to reduce the pain of blood glucose testing: ear versus thumb. *BMJ* 2000;321:20. doi:10.1136/bmj.321.7252.20
- 332 Martin GL, Trioux T, Gaudry S, Tubach F, Hajage D, Dechartres A. Association between lack of blinding and mortality results in critical care randomized controlled trials: a meta-epidemiological study. *Crit Care Med* 2021;49:1800-11. doi:10.1097/CCM.00000000000005065
- 333 Hróbjartsson A, Thomsen ASS, Emanuelsson F, et al. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ* 2012;344:e1119. doi:10.1136/bmj.e1119
- 334 Hróbjartsson A, Thomsen ASS, Emanuelsson F, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *CMAJ* 2013;185:E201-11. doi:10.1503/cmaj.120744
- 335 Hróbjartsson A, Thomsen ASS, Emanuelsson F, et al. Observer bias in randomized clinical trials with time-to-event outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *Int J Epidemiol* 2014;43:937-48. doi:10.1093/ije/dyt270
- 336 Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *Int J Epidemiol* 2014;43:1272-83. doi:10.1093/ije/dyu115
- 337 Bialy L, Vandermeer B, Lacaze-Masmonteil T, Dryden DM, Hartling L. A meta-epidemiological study to examine the association between bias and treatment effects in neonatal trials. *Evid Based Child Health* 2014;9:1052-9. doi:10.1002/ebch.1985
- 338 Armijo-Olivo S, Fuentes J, da Costa BR, Saltaji H, Ha C, Cummings GG. Blinding in physical therapy trials and its association with treatment effects: a meta-epidemiological study. *Am J Phys Med Rehabil* 2017;96:34-44. doi:10.1097/PHM.0000000000000521
- 339 Armijo-Olivo S, Dennett L, Arienti C, et al. Blinding in rehabilitation research: empirical evidence on the association between blinding and treatment effect estimates. *Am J Phys Med Rehabil* 2020;99:198-209. doi:10.1097/PHM.00000000000001377
- 340 Zeraatkar D, Pitre T, Diaz-Martinez JP, et al. Impact of allocation concealment and blinding in trials addressing treatments for COVID-19: a methods study. *Am J Epidemiol* 2023;192:1678-87. doi:10.1093/aje/kwad131
- 341 Moustgaard H, Clayton GL, Jones HE, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. *BMJ* 2020;368:l6802. doi:10.1136/bmj.l6802
- 342 Mouillet G, Efficace F, Thiery-Vuillemin A, et al. Investigating the impact of open label design on patient-reported outcome results in prostate cancer randomized controlled trials. *Cancer Med* 2020;9:7363-74. doi:10.1002/cam4.3335
- 343 Anthon CT, Granholm A, Perner A, Laake JH, Möller MH. No firm evidence that lack of blinding affects estimates of mortality in randomized clinical trials of intensive care interventions: a systematic review and meta-analysis. *J Clin Epidemiol* 2018;100:71-81. doi:10.1016/j.jclinepi.2018.04.016
- 344 Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. *Ann Intern Med* 2002;136:254-9. doi:10.7326/0003-4819-136-3-2002050002
- 345 Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ* 2000;321:504. doi:10.1136/bmj.321.7259.504
- 346 Boutron I, Estellat C, Guittet L, et al. Methods of blinding in reports of randomized controlled trials assessing pharmacologic treatments: a systematic review. *PLoS Med* 2006;3:e425. doi:10.1371/journal.pmed.0030425
- 347 Boutron I, Guittet L, Estellat C, Moher D, Hróbjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Med* 2007;4:e61. doi:10.1371/journal.pmed.0040061
- 348 Montri VM, Bhandari M, Devereaux PJ, Manns BJ, Ghali WA, Guyatt GH. In the dark: the reporting of blinding status in randomized controlled trials. *J Clin Epidemiol* 2002;55:787-90. doi:10.1016/S0895-4356(02)00446-8
- 349 Dechartres A, Trinquet L, Atal I, et al. Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study. *BMJ* 2017;357:j2490. doi:10.1136/bmj.j2490
- 350 Cheng K, Smyth RL, Motley J, O'Hea U, Ashby D. Randomized controlled trials in cystic fibrosis (1966-1997) categorized by time, design, and intervention. *Pediatr Pulmonol* 2000;29:1-7. doi:10.1002/(SICI)1099-0496(200001)29:1<1::AID-PPUL>3.0.CO;2-1
- 351 Penić A, Begić D, Balajić K, Kowalski M, Marušić A, Puljak L. Definitions of blinding in randomised controlled trials of interventions published in high-impact anaesthesiology journals: a methodological study and survey of authors. *BMJ Open* 2020;10:e035168. doi:10.1136/bmjjopen-2019-035168
- 352 Lang T. Masking or blinding? An unscientific survey of mostly medical journal editors on the great debate. *MedGenMed* 2000;2:E25.
- 353 Mills E, Prousky J, Raskin G, et al. The safety of over-the-counter niacin. A randomized placebo-controlled trial [ISRCTN18054903]. *BMJ Clin Pharmacol* 2003;3:4. doi:10.1186/1472-6904-3-4
- 354 Bello S, Moustgaard H, Hróbjartsson A. Unreported formal assessment of unblinding occurred in 4 of 10 randomized clinical trials, unreported loss of blinding in 1 of 10 trials. *J Clin Epidemiol* 2017;81:42-50. doi:10.1016/j.jclinepi.2016.08.002
- 355 Bello S, Moustgaard H, Hróbjartsson A. The risk of unblinding was infrequently and incompletely reported in 300 randomized clinical trial publications. *J Clin Epidemiol* 2014;67:1059-69. doi:10.1016/j.jclinepi.2014.05.007
- 356 Hróbjartsson A, Boutron I. Blinding in randomized clinical trials: imposed impartiality. *Clin Pharmacol Ther* 2011;90:732-6. doi:10.1038/clpt.2011.207
- 357 Schulz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. *BMJ* 1996;312:742-4. doi:10.1136/bmj.312.7033.742
- 358 Ferguson D, Glass KC, Waring D, Shapiro S. Turning a blind eye: the success of blinding reported in a random sample of randomised, placebo controlled trials. *BMJ* 2004;328:432. doi:10.1136/bmj.37952.631667.EE
- 359 Sackett DL. Turning a blind eye: why we don't test for blindness at the end of our trials. *BMJ* 2004;328:1136. doi:10.1136/bmj.328.7448.1136-a
- 360 Schulz KF, Altman DG, Moher D, Ferguson D. CONSORT 2010 changes and testing blindness in RCTs. *Lancet* 2010;375:1144-6. doi:10.1016/S0140-6736(10)60413-8
- 361 Freeman D, Emsley R, Diamond R, et al. Oxford Cognitive Approaches to Psychosis Trial Study Group. Comparison of a theoretically driven cognitive therapy (the Feeling Safe Programme) with befriending for the treatment of persistent persecutory delusions: a parallel, single-blind, randomised controlled trial. *Lancet Psychiatry* 2021;8:696-707. doi:10.1016/S2215-0366(21)00158-9

- 362 Wilby MJ, Best A, Wood E, et al. Surgical microdiscectomy versus transforaminal epidural steroid injection in patients with sciatica secondary to herniated lumbar disc (NERVES): a phase 3, multicentre, open-label, randomised controlled trial and economic evaluation. *Lancet Rheumatol* 2021;3:e347-56. doi:10.1016/S2665-9913(21)00036-9
- 363 Witham MD, Band M, Ahmed A, et al, BiCARB study group. Clinical and cost-effectiveness of oral sodium bicarbonate therapy for older patients with chronic kidney disease and low-grade acidosis (BiCARB): a pragmatic randomised, double-blind, placebo-controlled trial. *BMC Med* 2020;18:91. doi:10.1186/s12916-020-1542-9
- 364 Lang TA, Secic M. How to report statistics in medicine: annotated guidelines for authors, editors, and reviewers. Philadelphia: American College of Physicians; 1997.
- 365 Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. In: Altman DG, Machin D, Bryant TN, Gardner MJ, eds. *Statistics with Confidence: Confidence Intervals and Statistical Guidelines*. 2nd ed. BMJ Books, 2000: 171-90.
- 366 Berry DA. Interim analyses in clinical trials: classical vs. Bayesian approaches. *Stat Med* 1985;4:521-6. doi:10.1002/sim.4780040412
- 367 Bray R, Hartley A, Wenkert D, et al, Medical Outreach Team of the Drug Information Association Bayesian Scientific Working Group. Why are There not More Bayesian Clinical Trials? Ability to Interpret Bayesian and Conventional Statistics Among Medical Researchers. *Ther Innov Regul Sci* 2023;57:426-35. doi:10.1007/s43441-022-00482-1
- 368 Fors M, González P. Current status of Bayesian clinical trials for oncology, 2020. *Contemp Clin Trials Commun* 2020;20:100658. doi:10.1016/j.concrt.2020.100658
- 369 Lee JJ, Chu CT. Bayesian clinical trials in action. *Stat Med* 2012;31:2955-72. doi:10.1002/sim.5404
- 370 Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials* 2014;15:139. doi:10.1186/1745-6215-15-139
- 371 Sullivan TR, Morris TP, Kahan BC, Cuthbert AR, Yelland LN. Categorisation of continuous covariates for stratified randomisation: How should we adjust? *Stat Med* 2024;43:2083-95. doi:10.1002/sim.10060
- 372 Kahan BC, Rushton H, Morris TP, Daniel RM. A comparison of methods to adjust for continuous covariates in the analysis of randomised trials. *BMC Med Res Methodol* 2016;16:42. doi:10.1186/s12874-016-0141-3
- 373 Yu LM, Chan AW, Hopewell S, Deeks JJ, Altman DG. Reporting on covariate adjustment in randomised controlled trials before and after revision of the 2001 CONSORT statement: a literature review. *Trials* 2010;11:59. doi:10.1186/1745-6215-11-59
- 374 Saquib N, Saquib J, Ioannidis JP. Practices and impact of primary outcome adjustment in randomized controlled trials: meta-epidemiologic study. *BMJ* 2013;347:f4313. doi:10.1136/bmj.f4313
- 375 Guglieri M, Bushby K, McDermott MP, et al, FOR-DMD Investigators of the Muscle Study Group. Effect of different corticosteroid dosing regimens on clinical outcomes in boys with Duchenne muscular dystrophy: a randomized clinical trial. *JAMA* 2022;327:1456-68. doi:10.1001/jama.2022.4315
- 376 Davies M, Færch L, Jeppesen OK, et al, STEP 2 Study Group. Semaglutide 2·4 mg once a week in adults with overweight or obesity, and type 2 diabetes (STEP 2): a randomised, double-blind, double-dummy, placebo-controlled, phase 3 trial. *Lancet* 2021;397:971-84. doi:10.1016/S0140-6736(21)00213-0
- 377 André F, Hee Park Y, Kim SB, et al. Trastuzumab deruxtecan versus treatment of physician's choice in patients with HER2-positive metastatic breast cancer (DESTINY-Breast02): a randomised, open-label, multicentre, phase 3 trial. *Lancet* 2023;401:1773-85. doi:10.1016/S0140-6736(23)00725-0
- 378 Freeman J, Hendrie W, Jarrett L, et al. Assessment of a home-based standing frame programme in people with progressive multiple sclerosis (SUMS): a pragmatic, multi-centre, randomised, controlled trial and cost-effectiveness analysis. *Lancet Neurol* 2019;18:736-47. doi:10.1016/S1474-4422(19)30190-5
- 379 Nüesch E, Trelle S, Reichenbach S, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 2009;339:b3244. doi:10.1136/bmj.b3244
- 380 Chalmers I, Matthews R, Glasziou P, Boutron I, Armitage P. Trial analysis by treatment allocated or by treatment received? Origins of 'the intention-to-treat principle' to reduce allocation bias: Part 1. *J R Soc Med* 2023;116:343-50. doi:10.1177/01410768231203922
- 381 Chalmers I, Matthews R, Glasziou P, Boutron I, Armitage P. Trial analysis by treatment allocated or by treatment received? Origins of 'the intention-to-treat principle' to reduce allocation bias: Part 2. *J R Soc Med* 2023;116:386-94. doi:10.1177/01410768231203936
- 382 White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011;342:d40. doi:10.1136/bmj.d40
- 383 Lachin JM. Fallacies of last observation carried forward analyses. *Clin Trials* 2016;13:161-8. doi:10.1177/1740774515602688
- 384 Molnar FJ, Man-Son-Hing M, Hutton B, Ferguson DA. Have last-observation-carried-forward analyses caused us to favour more toxic dementia therapies over less toxic alternatives? A systematic review. *Open Med* 2009;3:e31-50.
- 385 White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. *Clin Trials* 2012;9:396-407. doi:10.1177/1740774512450098
- 386 Gaughran F, Stringer D, Wojewodzka G, et al. Effect of vitamin D supplementation on outcomes in people with early psychosis: the DFEND randomized clinical trial. *JAMA Netw Open* 2021;4:e2140858. doi:10.1001/jamanetworkopen.2021.40858
- 387 Choudry NK, Fifer S, Fontanet CP, et al, SPINE CARE Investigators. Effect of a biopsychosocial intervention or postural therapy on disability and health care spending among patients with acute and subacute spine pain: the SPINE CARE randomized clinical trial. *JAMA* 2022;328:2334-44. doi:10.1001/jama.2022.22625
- 388 Lewis GD, Voors AA, Cohen-Solal A, et al. Effect of omecamtiv mecarbil on exercise capacity in chronic heart failure with reduced ejection fraction: the METEORIC-HF randomized clinical trial. *JAMA* 2022;328:259-69. doi:10.1001/jama.2022.11016
- 389 Schlapbach LJ, Gibbons KS, Horton SB, et al, NITRIC Study Group, the Australian and New Zealand Intensive Care Society Clinical Trials Group (ANZICS CTG), and the ANZICS Paediatric Study Group (PSG). Effect of nitric oxide via cardiopulmonary bypass on ventilator-free days in young children undergoing congenital heart disease surgery: the NITRIC randomized clinical trial. *JAMA* 2022;328:38-47. doi:10.1001/jama.2022.9376
- 390 Alk EA, Shawwa K, Kahale LA, et al. Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide. *BMJ Open* 2015;5:e008431. doi:10.1136/bmjopen-2015-008431
- 391 Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomised controlled trials in major medical journals. *Clin Trials* 2004;1:368-76. doi:10.1191/1740774504cn032oa
- 392 Bell ML, Fiero M, Horton NJ, Hsu C-H. Handling missing data in RCTs: a review of the top medical journals. *BMC Med Res Methodol* 2014;14:118. doi:10.1186/1471-2288-14-118
- 393 Ibrahim F, Tom BDM, Scott DL, Prevost AT. A systematic review of randomised controlled trials in rheumatoid arthritis: the reporting and handling of missing data in composite outcomes. *Trials* 2016;17:272. doi:10.1186/s13063-016-1402-5
- 394 Joseph R, Sim J, Ogollah R, Lewis M. A systematic review finds variable use of the intention-to-treat principle in musculoskeletal randomised controlled trials with missing data. *J Clin Epidemiol* 2015;68:15-24. doi:10.1016/j.jclinepi.2014.09.002
- 395 Kahale LA, Diab B, Khamis AM, et al. Potentially missing data are considerably more frequent than definitely missing data: a methodological survey of 638 randomised controlled trials. *J Clin Epidemiol* 2019;106:18-31. doi:10.1016/j.jclinepi.2018.10.001
- 396 Kearney A, Rosala-Hallas A, Rainford N, et al. Increased transparency was required when reporting imputation of primary outcome data in clinical trials. *J Clin Epidemiol* 2022;146:60-7. doi:10.1016/j.jclinepi.2022.02.008
- 397 Khan NA, Torralba KD, Aslam F. Missing data in randomised controlled trials of rheumatoid arthritis drug therapy are substantial and handled inappropriately. *RMD Open* 2021;7:e001708. doi:10.1136/rmdopen-2021-001708
- 398 Tan P-T, Cro S, Van Vogt E, Szegedi M, Cornelius VR. A review of the use of controlled multiple imputation in randomised controlled trials with missing outcome data. *BMC Med Res Methodol* 2021;21:72. doi:10.1186/s12874-021-01261-6
- 399 Zhang Y, Flórez ID, Colunga Lozano LE, et al. A systematic survey on reporting and methods for handling missing participant data for continuous outcomes in randomized controlled trials. *J Clin Epidemiol* 2017;88:57-66. doi:10.1016/j.jclinepi.2017.05.017
- 400 Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3-15. doi:10.1177/096228029900800102
- 401 Kenward MG, Molenberghs G. Last observation carried forward: a crystal ball? *Biopharm Stat* 2009;19:872-88. doi:10.1080/10543400903105406
- 402 Morris TP, Kahan BC, White IR. Choosing sensitivity analyses for randomised trials: principles. *BMC Med Res Methodol* 2014;14:11. doi:10.1186/1471-2288-14-11
- 403 Food and Drug Administration. E9 (R1) statistical principles for clinical trials: addendum: estimands and sensitivity analysis in clinical trials. Guidance for Industry. May 2021.
- 404 Hussain JA, White IR, Johnson MJ, et al. Development of guidelines to reduce, handle and report missing data in palliative care trials: A multi-stakeholder modified nominal group technique. *Palliat Med* 2022;36:59-70. doi:10.1177/02692163211065597

- 405 Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393. doi:10.1136/bmj.b2393
- 406 Eisenberg MJ, Hébert-Losier A, Windle SB, et al, E3 Investigators. Effect of e-cigarettes plus counseling vs counseling alone on smoking cessation: a randomized clinical trial. *JAMA* 2020;324:1844-54. doi:10.1001/jama.2020.18889
- 407 Kehoe PG, Turner N, Howden B, et al, RADAR investigators. Safety and efficacy of losartan for the reduction of brain atrophy in clinically diagnosed Alzheimer's disease (the RADAR trial): a double-blind, randomised, placebo-controlled, phase 2 trial. *Lancet Neurol* 2021;20:895-906. doi:10.1016/S1474-4422(21)00263-5
- 408 Knight M, Chiocchia V, Partlett C, et al, ANODE collaborative group. Prophylactic antibiotics in the prevention of infection after operative vaginal delivery (ANODE): a multicentre randomised controlled trial. *Lancet* 2019;393:2395-403. doi:10.1016/S0140-6736(19)30773-1
- 409 Sandhu HK, Booth K, Furlan AD, et al. Reducing opioid use for chronic pain with a group-based intervention: a randomized clinical trial. *JAMA* 2023;329:1745-56. doi:10.1001/jama.2023.6454
- 410 Parpia S, Morris TP, Phillips MR, et al, Retina Evidence Trials InterNational Alliance (R.E.T.I.N.A.) Study Group. Sensitivity analysis in clinical trials: three criteria for a valid sensitivity analysis. *Eye (Lond)* 2022;36:2073-4. doi:10.1038/s41433-022-02108-0
- 411 Williamson SF, Grayling MJ, Mander AP, et al. Subgroup analyses in randomized controlled trials frequently categorized continuous subgroup information. *J Clin Epidemiol* 2022;150:72-9. doi:10.1016/j.jclinepi.2022.06.017
- 412 Sun X, Briel M, Busse JW, et al. The influence of study characteristics on reporting of subgroup analyses in randomised controlled trials: systematic review. *BMJ* 2011;342:d1569. doi:10.1136/bmj.d1569
- 413 Paratore C, Zichi C, Audisio M, et al. Subgroup analyses in randomized phase III trials of systemic treatments in patients with advanced solid tumours: a systematic review of trials published between 2017 and 2020. *ESMO Open* 2022;7:100593. doi:10.1016/j.esmoop.2022.100593
- 414 Brand KJ, Hafpelmeyer A, Haller B. A systematic review of subgroup analyses in randomised clinical trials in cardiovascular disease. *Clin Trials* 2021;18:351-60. doi:10.1177/1740774520984866
- 415 Brankovic M, Kardys I, Steyerberg EW, et al. Understanding of interaction (subgroup) analysis in clinical trials. *Eur J Clin Invest* 2019;49:e13145. doi:10.1111/eci.13145
- 416 Wallach JD, Sullivan PG, Trepanowski JF, Steyerberg EW, Ioannidis JP. Sex based subgroup differences in randomized controlled trials: empirical evidence from Cochrane meta-analyses. *BMJ* 2016;355:i5826. doi:10.1136/bmj.i5826
- 417 Liu P, Ioannidis JPA, Ross JS, et al. Age-treatment subgroup analyses in Cochrane intervention reviews: a meta-epidemiological study. *BMC Med* 2019;17:188. doi:10.1186/s12916-019-1420-8
- 418 Liu P, Ross JS, Ioannidis JP, Dhruva SS, Vasiliou V, Wallach JD. Prevalence and significance of race and ethnicity subgroup analyses in Cochrane intervention reviews. *Clin Trials* 2020;17:231-4. doi:10.1177/1740774519887148
- 419 Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006;332:1080. doi:10.1136/bmj.332.7549.1080
- 420 Royston P, Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. *Stat Med* 2004;23:2509-25. doi:10.1002/sim.1815
- 421 Beard DJ, Davies LJ, Cook JA, et al, TOPKAT Study Group. The clinical and cost-effectiveness of total versus partial knee replacement in patients with medial compartment osteoarthritis (TOPKAT): 5-year outcomes of a randomised controlled trial. *Lancet* 2019;394:746-56. doi:10.1016/S0140-6736(19)31281-4
- 422 Nguyen C, Boutron I, Baron G, et al. Intradiscal glucocorticoid injection for patients with chronic low back pain associated with active discopathy. *Ann Intern Med* 2017;166:547-56. doi:10.7326/M16-1700
- 423 Egger M, Jüni P, Bartlett C, CONSORT Group (Consolidated Standards of Reporting of Trials). Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001;285:1996-9. doi:10.1001/jama.285.15.1996
- 424 Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979;301:1410-2. doi:10.1056/NEJM197912273012602
- 425 May GS, DeMets DL, Friedman LM, Furberg C, Passamani E. The randomized clinical trial: bias in analysis. *Circulation* 1981;64:669-73. doi:10.1161/01.CIR.64.4.669
- 426 Altman DG, Cuzick J, Peto J. More on zidovudine in asymptomatic HIV infection. *N Engl J Med* 1994;330:1758-9. doi:10.1056/NEJM199406163302415
- 427 Meinert CL. Beyond CONSORT: need for improved reporting standards for clinical trials. *Consolidated Standards of Reporting Trials. JAMA* 1998;279:1487-9. doi:10.1001/jama.279.18.1487
- 428 Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med* 1987;317:426-32. doi:10.1056/NEJM198708133170706
- 429 LaCroix AZ, Ott SM, Ichikawa L, Scholes D, Barlow WE. Low-dose hydrochlorothiazide and preservation of bone mineral density in older adults. A randomized, double-blind, placebo-controlled trial. *Ann Intern Med* 2000;133:516-26. doi:10.7326/0003-4819-133-7-200010030-00010
- 430 Predel HG, Ebel-Boutin C, Lange R, Weiser T. A randomized, placebo- and active-controlled, multi-country, multi-center parallel group trial to evaluate the efficacy and safety of a fixed-dose combination of 400 mg ibuprofen and 100 mg caffeine compared with ibuprofen 400 mg and placebo in patients with acute lower back or neck pain. *J Pain Res* 2019;12:2771-83. doi:10.2147/JPR.S217045
- 431 Shuster JJ. Median follow-up in clinical trials. *J Clin Oncol* 1991;9:191-2. doi:10.1200/JCO.1991.9.1.191
- 432 Altman DG, De Stavola BL, Love SB, Stepniewska KA. Review of survival analyses published in cancer journals. *Br J Cancer* 1995;72:511-8. doi:10.1038/bjc.1995.364
- 433 Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, Purem A. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. *PLoS Med* 2005;2:e298. doi:10.1371/journal.pmed.0020298
- 434 Diggle L, Deeks J. Effect of needle length on incidence of local reactions to routine immunisation in infants aged 4 months: randomised controlled trial. *BMJ* 2000;321:931-3. doi:10.1136/bmj.321.7266.931
- 435 Pocock S, White I. Trials stopped early: too good to be true? *Lancet* 1999;353:943-4. doi:10.1016/S0140-6736(98)00379-1
- 436 Montori VM, Devereaux PJ, Adhikari NKJ, et al. Randomized trials stopped early for benefit: a systematic review. *JAMA* 2005;294:2203-9. doi:10.1001/jama.294.17.2203
- 437 Psaty BM, Rennie D. Stopping medical research to save money: a broken pact with researchers and patients. *JAMA* 2003;289:2128-31. doi:10.1001/jama.289.16.2128
- 438 Temple R. FDA perspective on trials with interim efficacy evaluations. *Stat Med* 2006;25:3245-9, discussion 3326-47. doi:10.1002/sim.2631
- 439 Fernandes RM, van der Lee JH, Offringa M. A systematic review of the reporting of Data Monitoring Committees' roles, interim analysis and early termination in pediatric clinical trials. *BMC Pediatr* 2009;9:77. doi:10.1186/1471-2431-9-77
- 440 Hahn J-Y, Song YB, Oh J-H, et al, SMART-CHOICE Investigators. Effect of P2Y12 inhibitor monotherapy vs dual antiplatelet therapy on cardiovascular events in patients undergoing percutaneous coronary intervention: the SMART-CHOICE randomized clinical trial. *JAMA* 2019;321:2428-37. doi:10.1001/jama.2019.8146
- 441 Liu M, Chen J, Wu Q, Zhu W, Zhou X. Adherence to the CONSORT statement and extension for nonpharmacological treatments in randomized controlled trials of bariatric surgery: A systematic survey. *Obes Rev* 2021;22:e13252. doi:10.1111/obr.13252
- 442 Gossec L, Tubach F, Dougadois M, Ravaud P. Reporting of adherence to medication in recent randomized controlled trials of 6 chronic diseases: a systematic literature review. *Am J Med Sci* 2007;334:248-54. doi:10.1097/MJG.0b013e318068d8de8
- 443 Zhang Z, Peluso MJ, Gross CP, Visconti CM, Kernan WN. Adherence reporting in randomized controlled trials. *Clin Trials* 2014;11:195-204. doi:10.1177/1740774513512565
- 444 Persch AC, Page SJ. Protocol development, treatment fidelity, adherence to treatment, and quality control. *Am J Occup Ther* 2013;67:146-53. doi:10.5014/ajot.2013.006213
- 445 Beets MW, von Klingraff L, Burkart S, et al. Impact of risk of generalizability biases in adult obesity interventions: A meta-epidemiological review and meta-analysis. *Obes Rev* 2022;23:e13369. doi:10.1111/obr.13369
- 446 Beets MW, Weaver RG, Ioannidis JPA, et al. Identification and evaluation of risk of generalizability biases in pilot versus efficacy/effectiveness trials: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2020;17:19. doi:10.1186/s12966-020-0918-y
- 447 Tang W, Cao Z, Han M, et al. Hydroxychloroquine in patients with mainly mild to moderate coronavirus disease 2019: open label, randomised controlled trial. *BMJ* 2020;369:m1849. doi:10.1136/bmj.m1849
- 448 Sackett DL. Clinician-trialist rounds: 5. Cointervention bias-how to diagnose it in their trial and prevent it in yours. *Clin Trials* 2011;8:440-2. doi:10.1177/1740774511410995

- 449 Higgins JPT, Altman DG, Gøtzsche PC, et al. Cochrane Bias Methods GroupCochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi:10.1136/bmj.d5928
- 450 Moutzouri E, Adam L, Feller M, et al. Low reporting of cointerventions in recent cardiovascular clinical trials: a systematic review. *J Am Heart Assoc* 2020;9:e014890. doi:10.1161/JAHA.119.014890
- 451 Bührer J, Del Giovane C, Gencer B, et al. Inadequate reporting of cointerventions, other methodological factors, and treatment estimates in cardiovascular trials: a meta-epidemiological study. *Mayo Clin Proc Innov Qual Outcomes* 2023;7:231-40. doi:10.1016/j.mayocpiqo.2023.04.010
- 452 Iudici M, Riveros C, Gabay C, Puéchal X, Berthelot JM, Ravaud P. Inadequate reporting of concomitant medications in rheumatology randomized controlled trials of pharmacologic interventions. *Semin Arthritis Rheum* 2019;49:453-8. doi:10.1016/j.smarthrit.2019.03.009
- 453 Lamb SE, Williamson EM, Heine PJ, et al. Strengthening and Stretching for Rheumatoid Arthritis of the Hand Trial (SARAH) Trial Team. Exercises to improve function of the rheumatoid hand (SARAH): a randomised controlled trial. *Lancet* 2015;385:421-9. doi:10.1016/S0140-6736(14)60998-3
- 454 Welch VA, Norheim OF, Jull J, Cookson R, Sommerfelt H, Tugwell P. CONSORT-Equity and Boston Equity Symposium. CONSORT-Equity 2017 extension and elaboration for better reporting of health equity in randomised trials. *BMJ* 2017;359:j5085. doi:10.1136/bmj.j5085
- 455 Senn SJ. Base logic: tests of baseline balance in randomized clinical trials. *Clin Res Regul Aff* 1995;12:171-82. doi:10.3109/10601339509019426
- 456 Austin PC, Manca A, Zwarenstein M, Juurlink DN, Stanbrook MB. A substantial and confusing variation exists in handling of baseline covariates in randomized controlled trials: a review of trials published in leading medical journals. *J Clin Epidemiol* 2010;63:142-53. doi:10.1016/j.jclinepi.2009.06.002
- 457 Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med* 2002;21:2917-30. doi:10.1002/sim.1296
- 458 de Boer MR, Waterlander WE, Kuijper LDJ, Steenhuis IHM, Twisk JWV. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. *Int J Behav Nutr Phys Act* 2015;12:4. doi:10.1186/s12966-015-0162-z
- 459 Altman DG. Comparability of randomised groups. *Statistician* 1985;34:125-36. doi:10.2307/2987510.
- 460 Beard DJ, Davies L, Cook JA, et al. ACL SNNAP Study Group. Rehabilitation versus surgical reconstruction for non-acute anterior cruciate ligament injury (ACL SNNAP): a pragmatic randomised controlled trial. *Lancet* 2022;400:605-15. doi:10.1016/S0140-6736(22)01424-6
- 461 Sola A, Kurlat I, Pollak A, et al. Early versus delayed neonatal administration of a synthetic surfactant--the judgment of OSIRIS. The OSIRIS Collaborative Group (open study of infants at high risk of or with respiratory insufficiency--the role of surfactant. *Lancet* 1992;340:1363-9. doi:10.1016/0140-6736(92)92557-V
- 462 Hussain JA, White IR, Langan D, et al. Missing data in randomized controlled trials testing palliative interventions pose a significant risk of bias and loss of power: a systematic review and meta-analyses. *J Clin Epidemiol* 2016;74:57-65. doi:10.1016/j.jclinepi.2015.12.003
- 463 Altman DG. Clinical trials and meta-analyses. In: Altman DG, Machin D, Bryant DM, Gardner MJ, eds. *Statistics with confidence*. 2nd ed. BMJ Books, 2000: 120-38.
- 464 International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Intern Med* 1997;126:36-47. doi:10.7326/0003-4819-126-1-199701010-00006
- 465 Altman DG. Confidence intervals in practice. In: Altman DG, Machin D, Bryant TN, Gardner MJ, eds. *Statistics with confidence*. 2nd ed. BMJ Books, 2000: 6-14.
- 466 Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *Br Med J (Clin Res Ed)* 1986;292:746-50. doi:10.1136/bmj.292.6522.746
- 467 Bailar JC3rd, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. *Ann Intern Med* 1988;108:266-73. doi:10.7326/0003-4819-108-2-266
- 468 Dwan K, Altman DG, Arnaiz JA, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One* 2008;3:e3081. doi:10.1371/journal.pone.0003081
- 469 Sorensen L, Gyrd-Hansen D, Kristiansen IS, Nexøe J, Nielsen JB. Laypersons' understanding of relative risk reductions: randomised cross-sectional study. *BMC Med Inform Decis Mak* 2008;8:31. doi:10.1186/1472-6947-8-31
- 470 Bobbio M, Demichelis B, Giustetto G. Completeness of reporting trial results: effect on physicians' willingness to prescribe. *Lancet* 1994;343:1209-11. doi:10.1016/S0140-6736(94)92407-4
- 471 Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 1992;117:916-21. doi:10.7326/0003-4819-117-11-916
- 472 Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;310:452-4. doi:10.1136/bmj.310.6977.452
- 473 Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ* 1999;319:1492-5. doi:10.1136/bmj.319.7223.1492
- 474 Scharf O, Colevas AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. *J Clin Oncol* 2006;24:3933-8. doi:10.1200/JCO.2005.05.3959
- 475 Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med* 2009;169:1756-61. doi:10.1001/archinternmed.2009.306
- 476 Williams MR, McKeown A, Pressman Z, et al. Adverse event reporting in clinical trials of intravenous and invasive pain treatments: an ACTTION systematic review. *J Pain* 2016;17:1137-49. doi:10.1016/j.jpain.2016.07.006
- 477 Mayo-Wilson E, Qureshi R, Hong H, Chen X, Li T. Harms were detected but not reported in six clinical trials of gabapentin. *J Clin Epidemiol* 2023;164:76-87. doi:10.1016/j.jclinepi.2023.10.014
- 478 Riveros C, Dechartres A, Perrodeau E, Haneef R, Boutron I, Ravaud P. Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. *PLoS Med* 2013;10:e1001566, discussion e1001566. doi:10.1371/journal.pmed.1001566
- 479 Jain A, Greig AVH, Jones A, et al. NINJA Collaborative. Effectiveness of nail bed repair in children with or without replacing the fingernail: NINJA multicentre randomized clinical trial. *Br J Surg* 2023;110:432-8. doi:10.1093/bjs/znad031
- 480 Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. *Science* 1977;198:679-84. doi:10.1126/science.333584
- 481 Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000;355:1064-9. doi:10.1016/S0140-6736(00)02039-0
- 482 Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992;116:78-84. doi:10.7326/0003-4819-116-1-78
- 483 Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 1991;266:93-8. doi:10.1001/jama.1991.03470010097038
- 484 Hahn S, Williamson PR, Hutton JL, Garner P, Flynn EV. Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. *Stat Med* 2000;19:3325-36. doi:10.1002/1097-0258(20001230)19:24<3325::AID-SIM827>3.0.CO;2-D
- 485 Bhandari M, Devereaux PJ, Li P, et al. Misuse of baseline comparison tests and subgroup analyses in surgical trials. *Clin Orthop Relat Res* 2006;447:247-51. doi:10.1097/01.blo.0000218736.23506.fe
- 486 Horton R. The rhetoric of research. *BMJ* 1995;310:985-7. doi:10.1136/bmj.310.6985.985
- 487 Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIN randomized controlled trial. *J Clin Oncol* 2014;32:4120-6. doi:10.1200/JCO.2014.56.7503
- 488 Checkett JX, Riddle J, Zaaza Z, Booze MA, Whitener JH, Vassar MB. An evaluation of spin in lower extremity joint trials. *J Arthroplasty* 2019;34:1008-12. doi:10.1016/j.arth.2019.01.016
- 489 Austin J, Smith C, Natarajan K, Som M, Wayant C, Vassar M. Evaluation of spin within abstracts in obesity randomized clinical trials: A cross-sectional review. *Clin Obes* 2019;9:e12292. doi:10.1111/cob.12292
- 490 Clarke M, Hopewell S, Chalmers I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report. *J R Soc Med* 2007;100:187-90. doi:10.1177/014107680710011415
- 491 Glasziou P, Jones M, Clarke M. Setting new research in the context of previous research: some options. *BMJ Evid Based Med* 2024;29:44-6. doi:10.1136/bmjebm-2023-112300
- 492 Hoffmann F, Allers K, Rombeke T, et al. Nearly 80 systematic reviews were published each day: Observational study on trends in epidemiology and reporting over the years 2000-2019. *J Clin Epidemiol* 2021;138:1-11. doi:10.1016/j.jclinepi.2021.05.022
- 493 Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev* 2009;2009:MR000006.
- 494 Aviand MS, Ioannidis JPA, Mashour GA. Independent discussion sections for improving inferential reproducibility in published research. *Br J Anaesth* 2019;122:413-20. doi:10.1016/j.bja.2018.12.010

- 495 Boden WE, O'Rourke RA, Teo KK, et al, COURAGE Trial Research Group. Optimal medical therapy with or without PCI for stable coronary disease. *N Engl J Med* 2007;356:1503-16. doi:10.1056/NEJMoa070829
- 496 Purcell GP, Donovan SL, Davidoff F. Changes to manuscripts during the editorial process: characterizing the evolution of a clinical paper. *JAMA* 1998;280:227-8. doi:10.1001/jama.280.3.227
- 497 Annals of Internal Medicine. Information for authors 2008 <https://www.annals.org/> [Accessed 15 January 2008].
- 498 Docherty M, Smith R. The case for structuring the discussion of scientific papers. *BMJ* 1999;318:1224-5. doi:10.1136/bmj.318.7193.1224
- 499 The BMJ. Article types and preparation. <https://www.bmj.com/about-bmj/resources-authors/article-types> [Accessed 18 January 2024]
- 500 Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ* 1995;311:485. doi:10.1136/bmj.311.7003.485
- 501 Rothwell PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials* 2006;1:e9. doi:10.1371/journal.pctr.0010009
- 502 Chauvin A, Ravaud P, Moher D, et al. Accuracy in detecting inadequate research reporting by early career peer reviewers using an online CONSORT-based peer-review tool (COBPeer) versus the usual peer-review process: a cross-sectional diagnostic study. *BMC Med* 2019;17:205. doi:10.1186/s12916-019-1436-0

**Web appendix 1:** CONSORT 2025 expanded checklist of detailed information to include when reporting a randomised trial