

# Guidelines for Reporting Outcomes in Trial Reports

## The CONSORT-Outcomes 2022 Extension

Nancy J. Butcher, PhD; Andrea Monsour, MPH; Emma J. Mew, MPH, MPhil; An-Wen Chan, MD, DPhil; David Moher, PhD; Evan Mayo-Wilson, DPhil; Caroline B. Terwee, PhD; Alyssandra Chee-A-Tow, MPH; Ami Baba, MRes; Frank Gavin, MA; Jeremy M. Grimshaw, MBChB, PhD; Lauren E. Kelly, PhD; Leena Saeed, BSc, BEd; Lehana Thabane, PhD; Lisa Askie, PhD; Maureen Smith, MEd; Mufiza Farid-Kapadia, MD, PhD; Paula R. Williamson, PhD; Peter Szatmari, MD; Peter Tugwell, MD; Robert M. Golub, MD; Suneeta Monga, MD; Sunita Vohra, MD; Susan Marlin, MSc; Wendy J. Ungar, MSc, PhD; Martin Offringa, MD, PhD

 Supplemental content

**IMPORTANCE** Clinicians, patients, and policy makers rely on published results from clinical trials to help make evidence-informed decisions. To critically evaluate and use trial results, readers require complete and transparent information regarding what was planned, done, and found. Specific and harmonized guidance as to what outcome-specific information should be reported in publications of clinical trials is needed to reduce deficient reporting practices that obscure issues with outcome selection, assessment, and analysis.

**OBJECTIVE** To develop harmonized, evidence- and consensus-based standards for reporting outcomes in clinical trial reports through integration with the Consolidated Standards of Reporting Trials (CONSORT) 2010 statement.

**EVIDENCE REVIEW** Using the Enhancing the Quality and Transparency of Health Research (EQUATOR) methodological framework, the CONSORT-Outcomes 2022 extension of the CONSORT 2010 statement was developed by (1) generation and evaluation of candidate outcome reporting items via consultation with experts and a scoping review of existing guidance for reporting trial outcomes (published within the 10 years prior to March 19, 2018) identified through expert solicitation, electronic database searches of MEDLINE and the Cochrane Methodology Register, gray literature searches, and reference list searches; (2) a 3-round international Delphi voting process (November 2018-February 2019) completed by 124 panelists from 22 countries to rate and identify additional items; and (3) an in-person consensus meeting (April 9-10, 2019) attended by 25 panelists to identify essential items for the reporting of outcomes in clinical trial reports.

**FINDINGS** The scoping review and consultation with experts identified 128 recommendations relevant to reporting outcomes in trial reports, the majority (83%) of which were not included in the CONSORT 2010 statement. All recommendations were consolidated into 64 items for Delphi voting; after the Delphi survey process, 30 items met criteria for further evaluation at the consensus meeting and possible inclusion in the CONSORT-Outcomes 2022 extension. The discussions during and after the consensus meeting yielded 17 items that elaborate on the CONSORT 2010 statement checklist items and are related to completely defining and justifying the trial outcomes, including how and when they were assessed (CONSORT 2010 statement checklist item 6a), defining and justifying the target difference between treatment groups during sample size calculations (CONSORT 2010 statement checklist item 7a), describing the statistical methods used to compare groups for the primary and secondary outcomes (CONSORT 2010 statement checklist item 12a), and describing the prespecified analyses and any outcome analyses not prespecified (CONSORT 2010 statement checklist item 18).

**CONCLUSIONS AND RELEVANCE** This CONSORT-Outcomes 2022 extension of the CONSORT 2010 statement provides 17 outcome-specific items that should be addressed in all published clinical trial reports and may help increase trial utility, replicability, and transparency and may minimize the risk of selective nonreporting of trial results.

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

**Corresponding Author:** Nancy J. Butcher, PhD, Peter Gilgan Centre for Research and Learning, The Hospital for Sick Children, 686 Bay St, Toronto, ON M5G 0A4, Canada (nancy.butcher@sickkids.ca).

JAMA. 2022;328(22):2252-2264. doi:10.1001/jama.2022.21022

**W**ell designed and properly conducted randomized clinical trials (RCTs) are the gold standard for producing primary evidence that informs evidence-based clinical decision-making. In RCTs, trial outcomes are used to assess the intervention effects on participants.<sup>1</sup> The Consolidated Standards of Reporting Trials (CONSORT) 2010 statement provided 25 reporting items for inclusion in published RCT reports.<sup>2,3</sup>

Fully reporting trial outcomes is important for replicating results, knowledge synthesis efforts, and preventing selective nonreporting of results. A scoping review revealed diverse and inconsistent recommendations on how to report trial outcomes in published reports by academic, regulatory, and other key sources.<sup>4</sup> Insufficient outcome reporting remains common across academic journals and disciplines; key information about outcome selection, definition, assessment, analysis, and changes from the prespecified outcomes (ie, from the trial protocol or the trial registry) is often poorly reported.<sup>5-9</sup> Such avoidable reporting issues have been shown to affect the conclusions drawn from systematic reviews and meta-analyses,<sup>10</sup> contributing to research waste.<sup>11</sup>

Although calls for improved reporting of trial outcomes have been made,<sup>5,12</sup> what constitutes useful, complete reporting of trial outcomes to knowledge users such as trialists, systematic reviewers, journal editors, clinicians, patients, and the public is unclear.<sup>4</sup> Two extensions (for harms in 2004 and for patient-reported outcomes in 2013)<sup>13,14</sup> of the CONSORT statement relevant to the reporting of specific types of trial outcomes exist; however, no standard reporting guideline for essential outcome-specific information applicable to all outcome types, populations, and trial designs is available.<sup>4</sup>

The aim of the CONSORT-Outcomes 2022 extension was to develop harmonized, evidence- and consensus-based outcome reporting standards for clinical trial reports.

## Methods

The CONSORT-Outcomes 2022 extension was developed as part of the Instrument for Reporting Planned Endpoints in Clinical Trials (InsPECT) project<sup>15</sup> in accordance with the Enhancing the Quality and Transparency of Health Research (EQUATOR) methodological framework for reporting guideline development.<sup>16</sup> Ethics approval was not required as determined by the research ethics committee at The Hospital for Sick Children. The development<sup>15</sup> of the CONSORT-Outcomes 2022 extension occurred in parallel with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)-Outcomes 2022 extension for clinical trial protocols.<sup>17</sup>

### Key Development Phases

First, we created an initial list of recommendations relevant to reporting outcomes for RCTs that were synthesized from consultation with experts and a scoping review of existing guidance for reporting trial outcomes (published within the 10 years prior to March 19, 2018) identified through expert solicitation, electronic database searches of MEDLINE and the Cochrane Methodology Register, gray literature searches, and reference list searches as described.<sup>4,18</sup> Second, a 3-round international Delphi voting process took place from November 2018 to February 2019 to identify

### Key Points

**Question** What outcome-specific information should be included in a published clinical trial report?

**Findings** Using an evidence-based and international consensus-based approach that applied methods from the Enhancing the Quality and Transparency of Health Research (EQUATOR) methodological framework, 17 outcome-specific reporting items were identified.

**Meaning** Inclusion of these items in clinical trial reports may enhance trial utility, replicability, and transparency and may help limit selective nonreporting of trial results.

additional items and assess the importance of each item, which was completed by 124 panelists from 22 countries (eTable 1 in the [Supplement](#)). Third, an in-person expert consensus meeting was held (April 9-10, 2019), which was attended by 25 panelists from 4 countries, including a patient partner and a public partner, to identify the set of essential items relevant to reporting outcomes for trial reports and establish dissemination activities. Selection and wording of the items was finalized at a postconsensus meeting by executive panel members and via email with consensus meeting panelists.

### Other Information

The detailed methods describing development of the CONSORT-Outcomes 2022 extension appear in eAppendix 1 in the [Supplement](#), including the number of items evaluated at each phase and the process toward the final set of included items (eFigure in the [Supplement](#)). The scoping review trial protocol and findings have been published<sup>4,18</sup> and appear in eAppendix 1 in the [Supplement](#) and the search strategy appears in eAppendix 2 in the [Supplement](#). The self-reported characteristics of the Delphi voting panelists and the consensus meeting panelists appear in eTables 1-2 in the [Supplement](#). Details regarding the patient and public partner involvement appear in eAppendix 1 in the [Supplement](#).

## Results

In addition to the inclusion of the CONSORT 2010 statement checklist items, the CONSORT-Outcomes 2022 extension recommends that a minimum of 17 outcome-specific reporting items be included in clinical trial reports, regardless of trial design or population. The scoping review and consultation with experts identified 128 recommendations relevant to reporting outcomes in trial reports, the majority (83%) of which were not included in the CONSORT 2010 statement. All recommendations were consolidated into 64 items for Delphi voting; after the Delphi survey process, 30 items met the criteria for further evaluation at the consensus meeting and possible inclusion in the CONSORT-Outcomes 2022 extension. The CONSORT 2010 statement checklist items and the 17 items added by the CONSORT-Outcomes 2022 extension appear in [Table 1](#).<sup>19</sup>

A fillable version of the checklist appears in eTables 3-4 in the [Supplement](#) and on the CONSORT website.<sup>20</sup> When using the updated checklist, users should refer to definitions of key

**Table 1. Recommended Checklist Items to Address in a Clinical Trial Report From the CONSORT 2010 Statement and the CONSORT-Outcomes 2022 Extension<sup>a</sup>**

Section	Item No.	CONSORT 2010 statement	Item No.	CONSORT-Outcomes 2022 Extension
<b>Title and abstract</b>				
Title and abstract	1a	Identification as a randomized trial in the title		
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)		
<b>Introduction</b>				
Background and objectives	2a	Scientific background and explanation of rationale		
	2b	Specific objectives or hypotheses		
<b>Methods</b>				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio		
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		
Participants	4a	Eligibility criteria for participants		
	4b	Settings and locations where the data were collected		
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered (for specific guidance see TIDieR checklist and guide) <sup>19</sup>		
Outcomes	6a	Completely defined prespecified primary and secondary outcome measures, including how and when they were assessed	6a.1	Provide a rationale for the selection of the domain for the trial's primary outcome
			6a.2	Describe the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, mean, proportion), and the time point for each outcome
			6a.3	If the analysis metric for the primary outcome represents within-participant change, define and justify the minimal important change in individuals
			6a.4	If the outcome data were continuous, but were analyzed as categorical (method of aggregation), specify the cutoff values used
			6a.5	If outcome assessments were performed at several time points after randomization, state the time points used for the analysis
			6a.6	If a composite outcome was used, define all individual components of the composite outcome
			6a.7	Identify any outcomes that were not prespecified in a trial registry or trial protocol
			6a.8	Provide a description of the study instruments used to assess the outcome (eg, questionnaires, laboratory tests) along with reliability, validity, and responsiveness in a population similar to the study sample
			6a.9	Describe who assessed the outcome (eg, nurse, parent) and any qualifications or trial-specific training necessary to administer the study instruments to assess the outcome
			6a.10	Describe any processes used to promote outcome data quality during data collection (eg, duplicate measurements) and after data collection (eg, range checks of outcome data values), or state where these details can be found
Sample size	7a	How sample size was determined	7a.1	Define and justify the target difference between treatment groups (eg, the minimal important difference)
	7b	When applicable, explanation of any interim analyses and stopping guidelines		
<b>Randomization</b>				
Sequence generation	8a	Method used to generate the random allocation sequence		
	8b	Type of randomization; details of any restriction (such as blocking and block size)		
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned		

(continued)

**Table 1. Recommended Checklist Items to Address in a Clinical Trial Report From the CONSORT 2010 Statement and the CONSORT-Outcomes 2022 Extension<sup>a</sup> (continued)**

Section	Item No.	CONSORT 2010 statement	Item No.	CONSORT-Outcomes 2022 Extension
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions		
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		
	11b	If relevant, description of the similarity of interventions		
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	12a.1	Describe any methods used to account for multiplicity in the analysis or interpretation of the primary and secondary outcomes (eg, coprimary outcomes, same outcome assessed at multiple time points, or subgroup analyses of an outcome)
			12a.2	State and justify any criteria for excluding any outcome data from the analysis and reporting, or report that no outcome data were excluded
			12a.3	Describe the methods used to assess patterns of missingness (eg, missing not at random), and describe the methods used to handle missing outcome items or entire assessments
			12a.4	Provide a definition of the outcome analysis population relating to nonadherence of the trial protocol (eg, as a randomized analysis)
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		
<b>Results</b>				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome		
	13b	For each group, losses and exclusions after randomization, together with reasons		
Recruitment	14a	Dates defining the periods of recruitment and follow-up		
	14b	Why the trial ended or was stopped		
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group		
Numbers analyzed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups		
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% CI)	17a.1	Include the results for all prespecified outcome analyses or state where the results can be found if not in this report
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory	18.1	If there were any analyses that were not prespecified, explain why they were performed
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms <sup>13</sup> )		
<b>Discussion</b>				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		
Generalizability	21	Generalizability (external validity, applicability) of the trial findings		
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		
<b>Other information</b>				
Registration	23	Registration number and name of trial registry		
Protocol	24	Where the full trial protocol can be accessed, if available		
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		

Abbreviations: CONSORT, Consolidated Standards of Reporting Trials; TIDieR, Template for Intervention Description and Replication.

<sup>a</sup> It is strongly recommended that this checklist be read in conjunction with the

CONSORT 2010 statement guidelines<sup>2,3</sup> for important clarification on the checklist items. The CONSORT 2010 statement checklist is distributed under the terms of the Creative Commons license.

## Box. Glossary of Terms Used in the CONSORT-Outcomes 2022 Extension

**Composite outcome:** A composite outcome consists of  $\geq 2$  component outcomes (eg, proportion of participants who died or experienced a nonfatal stroke). Participants who have experienced any of the events specified by the components are considered to have experienced the composite outcome.<sup>21,22</sup>

**CONSORT 2010:** Consolidated Standards of Reporting Trials (CONSORT) statement that was published in 2010.<sup>2,3</sup>

**CONSORT-Outcomes 2022 extension:** Additional essential checklist items describing outcome-related content that are not covered by the CONSORT 2010 statement.

**Construct validity:** The degree to which the scores reported in a trial are consistent with the hypotheses (eg, with regard to internal relationships, the relationships of the scores to other instruments, or relevant between-group differences) based on the assumption that the instrument validly measures the domain to be measured.<sup>30</sup>

**Content validity:** The degree to which the content of the study instrument is an adequate reflection of the domain to be measured.<sup>30</sup>

**Criterion validity:** The degree to which the scores of a study instrument are an adequate reflection of a gold standard.<sup>30</sup>

**Cross-cultural validity:** The degree to which the performance of the items on a translated or culturally adapted study instrument are an adequate reflection of the performance of the items using the original version of the instrument.<sup>30</sup>

**Minimal important change:** The smallest within-patient change that is considered important by patients, clinicians, or relevant others.<sup>4,5</sup>

The change may be in a score or unit of measure (continuous or ordinal measurements) or in frequency (dichotomous outcomes). This term is often used interchangeably in health literature with the term *minimal important difference*. In the CONSORT-Outcomes 2022 extension, the minimal important change conceptually refers to important intrapatient change (item 6a.3) and the minimal important difference refers to the important between-group difference. Minor variants of the term, such as minimum instead of minimal, or the addition of the adjective clinically or clinical are common (eg, the minimum clinically important change).<sup>23</sup>

**Minimal important difference:** The smallest between-group difference that is considered important by patients, clinicians, or relevant others.<sup>24-27</sup> The difference may be in a score or unit of measure (continuous or ordinal measurements) or in frequency (dichotomous outcomes). Minor variants of the term, such as minimum instead of minimal, or the addition of the adjective clinically or clinical are common (eg, the minimum clinically important difference).<sup>23</sup>

**Outcome:** Refers to what is being assessed to examine the effect of exposure to a health intervention.<sup>1</sup> The 5 core elements of a defined outcome appear in Table 2.

**Primary outcome:** The planned outcome that is most directly related to the primary objective of the trial.<sup>28</sup> It is typically the outcome used in the sample size calculation for trials with the primary objective of assessing efficacy or effectiveness.<sup>29</sup> Many trials have 1 primary outcome, but some have  $>1$ . The term *primary end point* is sometimes used in the medical literature when referring to the primary outcome.<sup>4</sup>

**Reliability:** The degree to which the measurement is free from error. Specifically, the extent to which scores have not changed for participants and are the same for repeated measures under several conditions (eg, using different sets of items from the same rating scale for internal consistency; over time or test-retest; by different persons on the same occasion or interrater; or by the same persons, such as raters or responders, on different occasions or intrarater).<sup>30</sup>

**Responsiveness:** The ability of a study instrument to accurately detect and measure change in the outcome domain over time.<sup>31,32</sup> Distinct from an instrument's construct validity and criterion validity, which refer to the validity of a single score, responsiveness refers to the validity of a change score (ie, longitudinal validity).<sup>30</sup>

**Secondary outcomes:** The outcomes prespecified in the trial protocol to assess any additional effects of the intervention.<sup>28</sup>

**Smallest worthwhile effect:** The smallest beneficial effect of an intervention that justifies the costs, potential harms, and inconvenience of the interventions as determined by patients.<sup>33</sup>

**SPIRIT 2013:** Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement that was published in 2013.<sup>35,36</sup>

**SPIRIT-Outcomes 2022 extension:** Additional essential checklist items describing outcome-related trial protocol content that are not covered by the SPIRIT 2013 statement.<sup>17</sup>

**Structural validity:** The degree to which the scores of a study instrument (eg, a patient questionnaire or a clinical rating scale) are an adequate reflection of the dimensionality of the domain to be measured.<sup>30</sup>

**Study instrument:** The scale or tool used to make an assessment. A study instrument may be a questionnaire, a clinical rating scale, a laboratory test, a score obtained through a physical examination or an observation of an image, or a response to a single question.<sup>34</sup>

**Target difference:** The value that is used in sample size calculations as the difference sought to be detected on the primary outcome between intervention groups and that should be considered realistic or important (such as the minimal important difference or the smallest worthwhile effect) by  $\geq 1$  key stakeholder groups.<sup>37,38</sup>

**Validity:** The degree to which a study instrument measures the domain it purports to measure.<sup>30</sup>

terms in the glossary<sup>21-38</sup> (Box) because variations in the terminology and definitions exist across disciplines and geographic areas. The 5 core elements of a defined outcome (with examples) appear in Table 2.<sup>39,40</sup>

Application of these new checklist items from the CONSORT-Outcomes 2022 extension, in conjunction with the CONSORT 2010 statement, ensures trial outcomes will be comprehensively defined and reported. The estimated list of key users, their proposed actions, and the consequential potential benefits of implementing the 17 CONSORT-Outcomes 2022 extension checklist items appears in eTable 5 in the Supplement and was generated from the con-

sensus meeting's knowledge translation session. Examination and application of these outcome reporting recommendations may be helpful for trial authors, journal editors, peer reviewers, systematic reviewers, patients, the public, and trial participants (eTable 5 in the Supplement).

This report contains a brief explanation of the 17 checklist items generated from the CONSORT-Outcomes 2022 extension. Guidance on how to report the existing checklist items can be found in the CONSORT 2010 statement,<sup>2</sup> in Table 1, and in an explanatory guideline report.<sup>41</sup> Additional items that may be useful to include in some trial reports or in associated trial documents

**Table 2. The 5 Core Elements of a Defined Outcome<sup>a</sup>**

Element No.	Element term	Definition used	Example 1	Example 2	Example 3
1	Domain <sup>b</sup>	Title or concept to describe ≥1 outcomes	Blood pressure	Depression	Death
2	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			
		<ul style="list-style-type: none"> <li>• Descriptive name</li> <li>• If applicable, the total score or the subscales that will be analyzed</li> </ul>	Systolic blood pressure measured with Omron upper arm blood pressure monitor	MADRS	All-cause mortality per the hospital database
3	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
4	Method of aggregation	The procedure for estimating the treatment effect			
		<ul style="list-style-type: none"> <li>• If the outcome will be treated as a continuous, categorical, or time-to-event variable</li> <li>• 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 cutoff values or categories compared</li> </ul>	Continuous variable	Binary variable	Time to event
5	Time point	The timing of follow-up measurements			
		<ul style="list-style-type: none"> <li>• When outcome measurements will be obtained</li> <li>• Which of the outcome measurements will be analyzed</li> </ul>	2, 4, and 12 wk after randomization	2, 4, 6, and 8 wk after randomization	Daily
			12 wk after randomization	8 wk after randomization	End of follow-up

Abbreviation: MADRS, Montgomery-Åsberg Depression Rating Scale.

<sup>a</sup> Content adapted from Zarin et al,<sup>39</sup> Mayo-Wilson et al,<sup>10</sup> and Chan et al.<sup>35</sup>

<sup>b</sup> An explicit and specific description of the outcome domain should be provided in the trial protocol, as appropriate, when defining the trial outcome. If an

outcome domain is broad, such as pain, a specific protocolized domain definition might be the daily average of the intensity of the sensation of pain expressed on a range from no pain to worst pain imaginable over a 24-hour window during an average day.<sup>40</sup>

(eg, the statistical analysis plan or in a clinical study report<sup>42</sup>) appear in eTable 6 in the Supplement, but were not considered essential reporting items for all trial reports.

**CONSORT-Outcomes 2022 Extension Checklist Items for the Descriptions of the Methods for Selecting, Measuring, and Describing the Outcomes Used in a Trial**  
**Item 6a.1. Provide a Rationale for the Selection of the Domain for the Trial's Primary Outcome**

This item expands on CONSORT 2010 statement checklist item 6a to explicitly ask for reporting on the rationale underlying the selection of the outcome domain for use as the primary outcome. At a broad conceptual level, the outcome's domain refers to the name or concept used to describe an outcome (eg, pain).<sup>10,39</sup> The word domain can be closely linked to and sometimes used equivalently with the terms *construct* and *attribute* in the literature.<sup>40</sup> Even though a complete outcome definition is expected to be provided in the trial report (as recommended by CONSORT 2010 statement checklist item 6a),<sup>2,40</sup> the rationale for the choice of the outcome domain for the trial's primary outcome is also essential to communicate because it underpins the purpose of the proposed trial.

Important aspects for the rationale may include (1) the importance of the outcome domain to the individuals involved in the trial (eg, patients, the public, clinicians, policy makers, funders, or health payers), (2) the expected effect of the intervention on the outcome domain, and (3) the ability to assess it accurately, safely, and feasibly during the trial. It also should be reported whether the selected outcome domain originated from a core outcome set

(ie, an agreed standardized set of outcomes that should be measured in all trials for a specific clinical area).<sup>43</sup>

**Item 6a.2. Describe the Specific Measurement Variable (eg, Systolic Blood Pressure), Analysis Metric (eg, Change From Baseline, Final Value, Time to Event), Method of Aggregation (eg, Mean, Proportion), and the Time Point for Each Outcome**

This item expands on CONSORT 2010 statement checklist item 6a that recommends completely defining prespecified primary and secondary outcomes, and provides specific recommendations that mirror SPIRIT 2013 statement checklist item 12 (and its explanatory text for defining trial outcomes).<sup>35</sup> CONSORT-Outcomes 2022 extension checklist item 6a.2 recommends describing each element of an outcome including its measurement variable, specific analysis metric, method of aggregation, and time point. Registers such as ClinicalTrials.gov already require that trials define their outcomes using this framework.<sup>10,35,39</sup> Failure to clearly report each element of the outcomes from a trial enables undetectable multiple testing, data cherry-picking, and selective nonreporting of results in the trial report compared with what was planned.<sup>10,44</sup>

**Item 6a.3. If the Analysis Metric for the Primary Outcome Represents Within-Participant Change, Define and Justify the Minimal Important Change in Individuals**

This item expands on CONSORT-Outcomes 2022 extension checklist item 6a.2. In cases in which the participant-level analysis metric for the primary outcome represents intraindividual change from an earlier value (such as those measured at baseline), a definition and

justification of what is considered the minimal important change (MIC) for the relevant study instrument should be provided. In the CONSORT-Outcomes 2022 extension, the MIC was defined as the smallest within-patient change that is considered important by patients, clinicians, or relevant others (common alternative terminologies appear in the Box).<sup>24,25,31</sup> The MIC is important to report for all trials that use a within-participant change metric, such as those that plan to analyze the proportion of participants showing a change larger than the MIC value in each treatment group (eg, to define the proportion who improved)<sup>45</sup> or in n-of-1 trial designs.<sup>46</sup>

Describing the MIC will facilitate understanding of the trial results and their clinical relevance by patients, clinicians, and policy makers. Users with trial knowledge may be interested in the MIC itself as a benchmark or, alternatively, in a value larger than the known MIC. Describing the justification for the selected MIC is important because there can be numerous MICs available for the same study instrument, with varying clinical relevance and methodological quality depending on how and in whom they were determined.<sup>47-50</sup> If the MIC is unknown for the study instrument with respect to the trial's population and setting, this should be reported.

**Item 6a.4. If the Outcome Data Were Continuous, but Were Analyzed as Categorical (Method of Aggregation), Specify the Cutoff Values Used**

This item expands on CONSORT-Outcomes 2022 extension checklist item 6a.2 to prompt authors, if applicable, to describe the prespecified cutoff values used to convert any outcome data collected on a continuous (or ordinal) scale into a categorical variable for their analyses.<sup>10,35</sup> Providing an explanation of the rationale for the choice of the cutoff value is recommended; it is not unusual for different trials to apply different cutoff values. The cutoff values selected are most useful when they have clear clinical relevance.<sup>51</sup> Reporting this information will help avoid multiple testing (known as "p-hacking"), data cherry-picking, and selective nonreporting of results in the trial report.<sup>10,44,52</sup>

**Item 6a.5. If Outcome Assessments Were Performed at Several Time Points After Randomization, State the Time Points Used for the Analysis**

This item expands on CONSORT-Outcomes 2022 extension checklist item 6a.2 regarding the time point to prompt authors, if applicable, to specify the time point used in the main analysis if outcome assessments were performed at multiple time points after randomization (eg, trial assessed blood pressure daily for 12 weeks after randomization). Specifying the preplanned time points of assessment used for the analyses will help limit the possibility of unplanned analyses of multiple assessment time points and the selective nonreporting of time points that did not yield large or significant results.<sup>35,39</sup>

Providing a rationale for the choice of time point is encouraged (eg, based on the expected clinical trajectory after the intervention or the duration of treatment needed to achieve a clinically meaningful exposure to treatment). The length of follow-up should be appropriate to the management decision the trial is designed to inform.<sup>53</sup>

**Item 6a.6. If a Composite Outcome Was Used, Define all Individual Components of the Composite Outcome**

A composite outcome consists of 2 or more component outcomes that may be related. Participants who have experienced any 1 of the

defined component outcomes comprising the composite outcome are considered to have experienced the composite outcome.<sup>21,22</sup> When used, composite outcomes should be prespecified, justified, and fully defined,<sup>51</sup> which includes a complete definition of each individual component outcome and a description of how those will be combined (eg, what analytic steps define the occurrence of the composite outcome).

However, composite outcomes can be difficult to interpret even when sufficiently reported. For example, a composite outcome can disguise treatment effects when the effects on the component outcomes go in opposite directions or when the component outcomes have different effect levels (eg, combining death and disability), furthering the need for quality reporting for every component.<sup>22,54</sup>

**Item 6a.7. Identify any Outcomes That Were not Prespecified in a Trial Registry or Trial Protocol**

Any outcomes that were not prespecified in the trial protocol or trial registry that were measured during the trial should be clearly identified and labeled. Outcomes that were not prespecified can result from the addition of an entirely new outcome domain that was not initially planned (eg, the unplanned inclusion and analysis of change in frequency of cardiovascular hospital admissions obtained from a hospital database). In addition, outcomes that differ from the prespecified outcomes in measurement variable, analysis metric, method of aggregation, and analysis time point are not prespecified. For example, if the trial reports on treatment success rates at 12 months instead of the prespecified time point of 6 months in the trial protocol, the 12-month rate should be identified as an outcome that was not prespecified with the specific change explained. Many reasons exist for changes in outcome data (eg, screening, diagnostic, and surveillance procedures may change; coding systems may change; or new adverse effect data may emerge). For fundamental changes to the primary outcome, investigators should report details (eg, the nature and timing of the change, motivation, whether the impetus arose from internal or external data sources, and who proposed and who approved these changes).

The addition of undeclared outcomes in trial reports is a major concern.<sup>7,55</sup> Among 67 trials published in 5 high-impact CONSORT-endorsing journals, there were 365 outcomes added (a mean of 5 undeclared outcomes per trial).<sup>7</sup> Less than 15% of the added outcomes were described as not being prespecified.<sup>7</sup> Determining whether reported outcomes match those in trial protocols or trial registries should not be left for readers to check for themselves, which is an onerous process (estimated review time of 1-7 hours per trial),<sup>7</sup> and is impossible in some cases (such as when the trial protocol is not publicly available). There can be good reasons to change study outcomes while a trial is ongoing, and authors should disclose these changes (CONSORT 2010 statement checklist item 6b) to avoid any potential appearance of reporting bias.

**Item 6a.8. Provide a Description of the Study Instruments Used to Assess the Outcome (eg, Questionnaires, Laboratory Tests) Along With Reliability, Validity, and Responsiveness in a Population Similar to the Study Sample**

The information needed to provide a sufficient description of the study instrument should be enough to allow replication of the trial

and interpretability of the results (eg, specify the version of the rating scale used, the mode of administration, and the make and model of the relevant laboratory instrument).<sup>35</sup> It is essential to summarize and provide references to empirical evidence that demonstrated sufficient reliability (eg, test-retest, interrater or intrarater reliability, and internal consistency), validity (eg, content, construct, criterion, cross-cultural, and structural validity), ability to detect change in the health outcome being assessed (ie, responsiveness) as appropriate for the type of study instrument, and enable comparison with a population similar to the study sample. Such evidence may be drawn from high-quality primary studies of measurement properties, from systematic reviews of measurement properties of study instruments, and from core outcome sets. Diagnostic test accuracy is relevant to report when the defined outcome relates to the presence or absence of a condition before and after treatment.<sup>56</sup>

CONSORT-Outcomes 2022 extension checklist item 6a.8 also recommends describing relevant measurement properties in a population similar to the study sample (or at least not substantially different) because measurement properties of study instruments cannot be presumed to be generalizable between different populations (eg, between different age groups).<sup>57</sup> If measurement properties of the study instrument were unknown for the population used, this can be stated with a rationale for why it was considered appropriate or necessary to use this instrument.

This information is critical to report because the quality and interpretation of the trial data rest on these measurement properties. For example, study instruments with poor content validity would not accurately reflect the domain that was intended to be measured, and study instruments with low interrater reliability would undermine the trial's statistical power<sup>35</sup> if an expected result was not discovered or accounted for in the planned power calculations.<sup>30-32</sup>

**Item 6a.9. Describe who Assessed the Outcome (eg, Nurse, Parent) and any Qualifications or Trial-Specific Training Necessary to Administer the Study Instruments to Assess the Outcome**

Substantially different responses, and therefore different trial results, can be obtained for many types of outcomes (eg, behavioral, psychological outcomes), depending on who is assessing the outcome of interest. This variability may result from differences in assessors' training or experience, different perspectives, or patient recall.<sup>58,59</sup> Assessments of a clinical outcome reported by a clinician, a patient, or a nonclinician observer or through a performance-based assessment are correspondingly classified by the US Food and Drug Administration as clinician-reported, patient-reported, observer-reported, and performance outcomes.<sup>60</sup>

For outcomes that could be assessed by various people, an explanation for the choice of outcome assessor made in the context of the trial should be provided. For outcomes that are not influenced by the outcome assessor (eg, plasma cholesterol levels), this information is less relevant. Professional qualifications or any trial-specific training necessary for trial personnel to function as outcome assessors is often relevant to describe<sup>35</sup> (eg, when using the second edition of the Wechsler Abbreviated Scale of Intelligence, an assessor with a PhD or PsyD and  $\geq 5$  years of experience with the relevant patient population and  $\geq 15$  prior administrations using this instrument or similar IQ assessments might be required). Details

regarding blinding of an assessor to the patient's treatment assignment and emerging trial results are covered in CONSORT 2010 statement checklist item 11a.

**Item 6a.10. Describe any Processes Used to Promote Outcome Data Quality During Data Collection (eg, Duplicate Measurements) and After Data Collection (eg, Range Checks of Outcome Data Values), or State Where These Details Can Be Found**

Providing a description of any of the processes used to promote outcome data quality during and after data collection in a trial provides transparency and facilitates appraisal of the quality of the trial's data. For example, subjective outcome assessments may have been performed in duplicate (eg, pathology assessments) or a central adjudication committee may have been used to ensure independent and accurate outcome assessments. Other common examples include verifying the data are in the proper format (eg, integer), the data are within an expected range of values, and the data are reviewed with independent source document verification (eg, by an external trial monitor).<sup>35</sup> The trial report should include a full description or a brief summary with reference to where the complete information can be found (eg, an open access trial protocol).

**CONSORT-Outcomes 2022 Extension Checklist Item for the Description of the Sample Size**

**Item 7a.1. Define and Justify the Target Difference Between Treatment Groups (eg, the Minimal Important Difference)**

This item expands on CONSORT 2010 statement checklist item 7a for reporting how sample size was determined to prompt authors to report the target difference used to inform the trial's sample size calculation. The target difference is the value used in sample size calculations as the difference sought to be detected in the primary outcome between the intervention groups at the specific time point for the analysis that should be considered realistic or important by 1 or more key stakeholder groups.<sup>37</sup> The Difference Elicitation in Trials project has published extensive evidence-based guidance on selecting a target difference for a trial, sample size calculation, and reporting.<sup>37,38</sup> The target difference may be the minimal important difference (MID; the smallest difference between patients perceived as important)<sup>24,26,27</sup> or the smallest worthwhile effect (the smallest beneficial effect of an intervention that justifies the costs, harms, and inconvenience of the interventions as determined by patients).<sup>33</sup> Because there can be different pragmatic or clinical factors informing the selected target difference (eg, the availability of a credible MID for the study instrument used to assess the primary outcome),<sup>47</sup> and numerous different options available (eg, 1 of several MIDs or values based on pilot studies),<sup>47</sup> it is important to explain why the chosen target difference was selected.<sup>23,48,49</sup>

**CONSORT-Outcomes 2022 Extension Checklist Items for the Descriptions of the Statistical Methods**

**Item 12a.1. Describe any Methods Used to Account for Multiplicity in the Analysis or Interpretation of the Primary and Secondary Outcomes (eg, Coprimary Outcomes, Same Outcome Assessed at Multiple Time Points, or Subgroup Analyses of an Outcome)**

This item extends CONSORT 2010 statement checklist item 12a to prompt authors to describe any statistical methods used to



account for multiplicity relating to the analysis or interpretation of the results. Outcome multiplicity issues are common in trials and deserve particular attention when there are coprimary outcomes, multiple possible time points resulting from the repeated assessment of a single outcome, multiple planned analyses of a single outcome (eg, interim or subgroup analysis, multigroup trials), or numerous secondary outcomes for analysis.<sup>61</sup>

The methods used to account for such forms of multiplicity include statistical methods (eg, family-wise error rate approaches) or descriptive approaches (eg, noting that the analyses are exploratory, placing the results in the context of the expected number of false-positive outcomes).<sup>61,62</sup> Such information may be briefly described in the text of the report or described in more detail in the statistical analysis plan.<sup>63</sup> Authors may report if no methods were used to account for multiplicity (eg, not applicable or were not considered necessary).

#### **Item 12a.2. State and Justify any Criteria for Excluding any Outcome Data From the Analysis and Reporting, or Report That no Outcome Data Were Excluded**

This item extends CONSORT 2010 statement checklist item 12a to recommend that authors (1) state and justify any criteria applied for excluding certain outcome data from the analysis or (2) report that no outcome data were excluded. This is in reference to explicitly and intentionally excluded outcome data, such as in the instance of too many missing items from a participant's completed questionnaire, or through other well-justified exclusion of outliers for a particular outcome. This helps the reader to interpret the reported results. This information may be presented in the CONSORT flow diagram where the reasons for outcome data exclusion are stated for each outcome by treatment group.

#### **Item 12a.3. Describe the Methods Used to Assess Patterns of Missingness (eg, Missing not at Random), and Describe the Methods Used to Handle Missing Outcome Items or Entire Assessments**

The occurrence of missing participant outcome data in trials is common, and in many cases, this missingness is not random, meaning it is related to either allocation to a treatment group, patient-specific (prognostic) factors, or the occurrence of a specific health outcome.<sup>64,65</sup> When there are missing data, CONSORT-Outcomes 2022 extension checklist item 12a.3 recommends describing (1) any methods used to assess or identify the pattern of missingness and (2) any methods used to handle missing outcomes or entire assessments (the choice of which is informed by the identified pattern of missingness) in the statistical analysis (eg, multiple imputation, complete case, based on likelihood, and inverse-probability weighting).

It is critical to provide information about the patterns and handling of any missing data because missing data can lead to reduced power of the trial, affect its conclusions, and affect whether trials are at low or high risk of bias, depending on the pattern of missingness.<sup>66,67</sup> A lack of clarity about the magnitude of the missingness and about how missing data were handled in the analysis makes it impossible for meta-analysts to accurately extract sample sizes needed to weight studies in their pooled estimates and prevents accurate assessment of any risk of bias arising from missing data in the reported results.<sup>67,68</sup> This checklist item is not

applicable if there is a complete data set, and it may be unimportant if the amount of missing data can be considered negligible.

Patterns of missingness (also referred to as missing data mechanisms) include missing completely at random, missing at random, and missing not at random and require description in trials to help readers and meta-analysts determine which patterns are present in data sets.<sup>69</sup> Some of the missing data may still be able to be measured (eg, via concerted follow-up efforts with a subset of the trial participants with missing data) to help distinguish between missing at random and missing not at random.<sup>70</sup> The pattern of missingness relates to the choice of the methods used to handle missing outcomes or entire assessments (eg, multiple imputation and maximum likelihood analyses assume the data are at least missing at random) and is essential to report. Any sensitivity analyses that were conducted to assess the robustness of the trial results (eg, using different methods to handle missing data) should be reported.<sup>35,64</sup>

#### **Item 12a.4. Provide a Definition of the Outcome Analysis Population Relating to Nonadherence of the Trial Protocol (eg, as a Randomized Analysis)**

Trial outcome data can be analyzed in many ways that can lead to different results. The general reporting principle described in the CONSORT 2010 statement was to "describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results."<sup>2</sup> Item 12a.4 extends CONSORT 2010 statement checklist item 12a to recommend that authors report the definition of the outcome analysis population used for each analysis as it relates to nonadherence of the trial protocol. For each analysis, information on whether the investigators included all participants who were randomized to the group to which they were originally allocated (ie, intention-to-treat analysis) has been widely recognized to be particularly important to the critical appraisal and interpretation of trial findings.<sup>2,35</sup>

Because amounts of missing data may vary among different outcomes and the reasons data are missing may also vary, CONSORT-Outcomes 2022 extension checklist item 12a.4 specifies reporting the definition of the outcome analysis population used in the statistical analyses. For example, a complete data set may be available to analyze the outcome of mortality but not for patient-reported outcomes within the same trial. In another example, analysis of harms might be restricted to participants who received the trial intervention so the absence or occurrence of harm was not attributed to a treatment that was never received.<sup>35</sup>

#### **CONSORT-Outcomes 2022 Extension Checklist Item for the Description of Outcomes and Estimation**

##### **Item 17a.1. Include the Results for all Prespecified Outcome Analyses or State Where the Results Can Be Found if not in This Report**

This item expands on CONSORT 2010 statement checklist item 17a on outcomes and estimation to remind authors to ensure that they have reported the results for all outcome analyses that were prespecified in the trial protocol or statistical analysis plan.<sup>68</sup> Although this is expected to be standard practice,<sup>2</sup> the information available in the trial report is often insufficient regarding prespecified analyses for the reader to determine whether there was selective nonreporting of any trial results.<sup>71</sup> When it is not feasible to report on all prespecified analyses in a single trial report (eg, trials with a large

number of prespecified secondary outcomes), authors should report where the results of any other prespecified outcome analyses can be found (eg, in linked publications or an online repository) or signal their intention to report later in the case of longer-term follow-up.

### CONSORT-Outcomes 2022 Extension Checklist Item for the Description of Ancillary Analyses

#### Item 18.1. If There Were any Analyses That Were not Prespecified, Explain why They Were Performed

A recent study of adherence showed that prespecified statistical analyses remain low in published trials, with unexplained discrepancies between the prespecified and reported analyses.<sup>71</sup> This item extends CONSORT 2010 statement checklist item 18 on ancillary analyses to recommend that an explanation should be provided for any analyses that were not prespecified (eg, in the trial protocol or statistical analysis plan), but that are being reported in the trial report. These types of analyses can be called either exploratory analyses or analyses that were not prespecified. Communicating the rationale for any analyses that were not prespecified, but which were performed and reported, is important for trial transparency and for correct appraisal of the trial's credibility. It can be important to state when such additional analyses were performed (eg, before or after seeing any results from comparative analyses for other outcomes). Multiple analyses of the same data create a risk for false-positive findings and selective reporting of analyses that were not prespecified could lead to bias.

## Discussion

The CONSORT-Outcomes 2022 extension provides evidence- and consensus-based guidance for reporting outcomes in published clinical trial reports, extending the CONSORT 2010 statement checklist with 17 additional reporting items and harmonizing reporting recommendations with guidance from the SPIRIT-Outcomes 2022 extension.<sup>17</sup> Alignment across these 2 extension guidelines creates a cohesive continuum of reporting from the trial protocol to the completed trial that will facilitate both the researcher's production of the trial protocol and trial report and, importantly, any assessment of the final report's adherence to the trial protocol.<sup>20</sup> Similar to the CONSORT 2010 statement,<sup>41</sup> the CONSORT-Outcomes 2022 extension applies to the content of the trial report, including the tables and figures and online-only supplementary material.<sup>72</sup> The current recommendations are similarly not prescriptive regarding the structure or location of reporting this information; authors should "address checklist items somewhere in the article, with ample detail and lucidity."<sup>41</sup>

Users of the CONSORT-Outcomes 2022 extension checklist should note that these additional checklist items represent the minimum essential items for outcomes reporting and are being added to the CONSORT 2010 statement<sup>2,3</sup> guidelines to maximize trial utility, transparency, replication, and limit selective nonreporting of results (eTable 5 in the [Supplement](#)). In some cases, it may be important to report additional outcome-specific information in trial reports,<sup>4</sup> such as those in eTable 6 in the [Supplement](#) or refer to CONSORT-PRO for PRO-specific reporting guidance<sup>14</sup> and the CONSORT extension for reporting harms.<sup>13</sup> Authors adhering to

the CONSORT-Outcomes 2022 extension should explain why any item is not relevant to their trial. For example, this extension checklist, which is for reporting systematically assessed outcomes, might not be applicable to outcomes that are not systematically collected or prespecified such as spontaneously reported adverse events. When constrained by journal word count, authors can refer to open access trial protocols, statistical analysis plans, trial registry data, or provide online-only supplementary materials.

We anticipate that the key users of the CONSORT-Outcomes 2022 extension will be trial authors, journal editors, peer reviewers, systematic reviewers, meta-analysis researchers, academic institutions, patients (including trial participants), and the broader public (eTable 5 in the [Supplement](#)). Use of this extension by these groups may help improve trial utility, transparency, and replication. Patient and public engagement was successfully embedded into a consensus meeting for a methodologically complex topic, a rarity in reporting guideline development to date. Future reporting guideline development should engage patients and members of the public throughout the process. The CONSORT-Outcomes 2022 extension will be disseminated as outlined previously,<sup>15</sup> including through the EQUATOR Network and the CONSORT website. End users can provide their input on the content, clarity, and usability online,<sup>73</sup> which will inform any future updates.

### Limitations

This study has several limitations. First, the included checklist items are appropriate for systematically collected outcomes, including most potential benefits and some harms; however, other items might be applicable for reporting harms not systematically assessed.<sup>74</sup>

Second, because these checklist items are not yet integrated in the main CONSORT checklist, finding and using multiple checklists may be considered burdensome by some authors and editors, which may affect uptake.<sup>75</sup> Future efforts to integrate these additional checklist items in the main CONSORT checklist might promote implementation in practice.

Third, although a large, multinational group of experts and end users was involved in the development of these recommendations with the aim of increasing usability among the broader research community, the Delphi voting results could have been affected by a non-response bias because panelists were self-selecting (ie, interested individuals signed up to take part in the Delphi voting process).

Fourth, the consensus meeting panelists were purposively sampled based on their expertise and roles relevant to clinical trial conduct, oversight, and reporting.<sup>15</sup> The views of individuals not well represented by the consensus meeting panelists (eg, trialists outside North America and Europe) may differ. The systematic and evidence-based approach<sup>15,16</sup> used to develop this guideline, including a rigorous scoping review of outcome reporting guidance,<sup>4,18</sup> will help mitigate the potential effect of these limitations.

## Conclusions

This CONSORT-Outcomes 2022 extension of the CONSORT 2010 statement provides 17 outcome-specific items that should be addressed in all published clinical trial reports and may help increase trial utility, replicability, and transparency and may minimize the risk of selective nonreporting of trial results.

## ARTICLE INFORMATION

**Accepted for Publication:** October 25, 2022.

**Author Affiliations:** Child Health Evaluative Sciences, The Hospital for Sick Children Research Institute, Toronto, Ontario, Canada (Butcher, Monsour, Mew, Chee-A-Tow, Baba, Saeed, Farid-Kapadia, Ungar, Offringa); Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada (Butcher, Monga); Department of Chronic Disease Epidemiology, School of Public Health, Yale University, New Haven, Connecticut (Mew); Department of Medicine, Women's College Research Institute, University of Toronto, Toronto, Ontario, Canada (Chan); Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada (Moher); School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada (Moher, Tugwell); Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill (Mayo-Wilson); Amsterdam University Medical Centers, Vrije Universiteit, Department of Epidemiology and Data Science, Amsterdam, the Netherlands (Terwee); Department of Methodology, Amsterdam Public Health Research Institute, Amsterdam, the Netherlands (Terwee); public panel member, Toronto, Ontario, Canada (Gavin); Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada (Grimshaw); Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada (Grimshaw, Tugwell); Department of Pharmacology and Therapeutics, University of Manitoba, Winnipeg, Canada (Kelly); Children's Hospital Research Institute of Manitoba, Winnipeg, Canada (Kelly); Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada (Thabane); NHMRC Clinical Trials Centre, University of Sydney, Sydney, New South Wales, Australia (Askie); patient panel member, Ottawa, Ontario, Canada (Smith); MRC-NIHR Trials Methodology Research Partnership, Department of Health Data Science, University of Liverpool, Liverpool, England (Williamson); Cundill Centre for Child and Youth Depression, Centre for Addiction and Mental Health, Toronto, Ontario, Canada (Szatmari); Department of Psychiatry, The Hospital for Sick Children, Toronto, Ontario, Canada (Szatmari, Monga); Bruyère Research Institute, Ottawa, Ontario, Canada (Tugwell); Ottawa Hospital Research Institute, Ottawa, Ontario, Canada (Tugwell); Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois (Golub); Departments of Pediatrics and Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Canada (Vohra); Clinical Trials Ontario, Toronto, Canada (Marlin); Department of Public Health Sciences, Queen's University, Kingston, Ontario, Canada (Marlin); Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, Ontario, Canada (Ungar, Offringa); Division of Neonatology, The Hospital for Sick Children, Toronto, Ontario, Canada (Offringa).

**Author Contributions:** Dr Butcher had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.  
**Concept and design:** All authors.

**Acquisition, analysis, or interpretation of data:** All authors.

**Drafting of the manuscript:** Butcher.

**Critical revision of the manuscript for important intellectual content:** All authors.

**Statistical analysis:** Monsour.

**Obtained funding:** Offringa.

**Administrative, technical, or material support:** Monsour, Mew, Baba.

**Supervision:** Butcher, Offringa.

**Conflict of Interest Disclosures:** Dr Butcher reported receiving grant funding from CHILD-BRIGHT and the Cundill Centre for Child and Youth Depression at the Centre for Addiction and Mental Health (CAMH) and receiving personal fees from Nobias Therapeutics. Ms Mew reported receiving salary support through a Canadian Institutes of Health Research doctoral foreign study award. Dr Kelly reported receiving funding from the Canadian Cancer Society, Research Manitoba, the Children's Hospital Research Institute of Manitoba, Mitacs, and the SickKids Foundation. Dr Askie reported being a co-convenor of the Cochrane Prospective Meta-Analysis Methods Group. Dr Farid-Kapadia reported currently being an employee of Hoffmann La-Roche and holding shares in the company. Dr Williamson reported chairing the COMET initiative management group. Dr Szatmari reported receiving funding from the CAMH. Dr Tugwell reported co-chairing the OMERACT executive committee; receiving personal fees from the Reformulary Group, UCB Pharma GmbH, Parexel International, PRA Health Sciences, Amgen, AstraZeneca, Bristol Myers Squibb, Celgene, Eli Lilly, Genentech, Roche, Genzyme, Sanofi, Horizon Therapeutics, Merck, Novartis, Pfizer, PPD Inc, QuintilesIMS (now IQVIA), Regeneron Pharmaceuticals, Savient Pharmaceuticals, Takeda Pharmaceutical Co Ltd, Vertex Pharmaceuticals, Forest Pharmaceuticals, and Bioiberica; serving on data and safety monitoring boards for UCB Pharma GmbH, Parexel International, and PRA Health Sciences; and receiving unrestricted educational grants from the American College of Rheumatology and the European League of Rheumatology. Dr Monga reported receiving funding from the Cundill Centre for Child and Youth Depression at CAMH; receiving royalties from Springer for *Assessing and Treating Anxiety Disorders in Young Children*; and receiving personal fees from the TD Bank Financial Group for serving as a chair in child and adolescent psychiatry. Dr Ungar reported being supported by the Canada research chair in economic evaluation and technology assessment in child health. No other disclosures were reported.

**Funding/Support:** This work, project 148953, received financial support from the Canadian Institutes of Health Research (supported the work of Drs Butcher, Kelly, Szatmari, Monga, and Offringa and Mss Saeed and Marlin).

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

**Disclaimers:** Dr Golub is Executive Deputy Editor of *JAMA*, but he was not involved in any of the

decisions regarding review of the manuscript or its acceptance. This article reflects the views of the authors, the Delphi panelists, and the consensus meeting panelists and may not represent the views of the broader stakeholder groups, the authors' institutions, or other affiliations.

**Additional Contributions:** We gratefully acknowledge the additional contributions made by the project core team, the executive team, the operations team, the Delphi panelists, and the international consensus meeting panelists (eAppendix 3 in the Supplement). We thank Andrea Chiaramida, BA, for administrative project support and Lisa Stallwood, MSc, for administrative manuscript support (both with The Hospital for Sick Children). We thank Petros Pechlivanoglou, PhD, and Robin Hayeems, PhD, for piloting and providing feedback on the format of the Delphi survey (both with The Hospital for Sick Children). We thank Roger F. Soll, MD (Cochrane Neonatal and the Division of Neonatal-Perinatal Medicine, Larner College of Medicine, University of Vermont), and James Webbe, MB BChir, PhD, and Chris Gale, MBBS, PhD (both with Neonatal Medicine, School of Public Health, Imperial College London), for pilot testing the CONSORT-Outcomes 2022 extension checklist. None of these individuals received compensation for their role in the study.

**Additional Information:** The project materials and data are publicly available on the Open Science Framework at <https://osf.io/arwy8/>.

## REFERENCES

1. COSMIN. Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) is an initiative of an international multidisciplinary team of researchers. Accessed April 13, 2021. <https://www.cosmin.nl/>
2. 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
3. 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
4. Butcher NJ, Mew EJ, Monsour A, Chan A-W, Moher D, Offringa M. Outcome reporting recommendations for clinical trial protocols and reports: a scoping review. *Trials*. 2020;21(1):620. doi:10.1186/s13063-020-04440-w
5. Bhaloo Z, Adams D, Liu Y, et al. Primary Outcomes Reporting in Trials (PORTal): a systematic review of inadequate reporting in pediatric randomized controlled trials. *J Clin Epidemiol*. 2017; 81:33-41. doi:10.1016/j.jclinepi.2016.09.003
6. Chan AW, Pello A, Kitchen J, et al. Association of trial registration with reporting of primary outcomes in protocols and publications. *JAMA*. 2017;318(17):1709-1711. doi:10.1001/jama.2017.13001
7. 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(1):118. doi:10.1186/s13063-019-3173-2
8. Monsour A, Mew EJ, Patel S, et al. Primary outcome reporting in adolescent depression clinical trials needs standardization. *BMC Med Res Methodol*. 2020;20(1):129. doi:10.1186/s12874-020-01019-6

9. Kampman JM, Sperna Weiland NH, Hollmann MW, Repping S, Hermanides J. High incidence of outcome switching observed in follow-up publications of randomized controlled trials: meta-research study. *J Clin Epidemiol*. 2021;137:236-240. doi:10.1016/j.jclinepi.2021.05.003
10. Mayo-Wilson E, Fusco N, Li T, Hong H, Canner JK, Dickersin K; MUDS investigators. Multiple outcomes and analyses in clinical trials create challenges for interpretation and research synthesis. *J Clin Epidemiol*. 2017;86:39-50. doi:10.1016/j.jclinepi.2017.05.007
11. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*. 2014;383(9913):267-276. doi:10.1016/S0140-6736(13)62228-X
12. Monga S, Offringa M, Butcher NJ, Szatmari P. From research to practice: the importance of appropriate outcome selection, measurement, and reporting in pediatric mental health research. *J Am Acad Child Adolesc Psychiatry*. 2020;59(4):497-500. doi:10.1016/j.jaac.2019.08.468
13. Ioannidis JP, Evans SJ, Gøtzsche PC, et al; CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med*. 2004;141(10):781-788. doi:10.7326/0003-4819-141-10-200411160-00009
14. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD; CONSORT PRO Group. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA*. 2013;309(8):814-822. doi:10.1001/jama.2013.879
15. Butcher NJ, Monsour A, Mew EJ, et al. Improving outcome reporting in clinical trial reports and protocols: study protocol for the Instrument for reporting Planned Endpoints in Clinical Trials (InsPECT). *Trials*. 2019;20(1):161. doi:10.1186/s13063-019-3248-0
16. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med*. 2010;7(2):e1000217. doi:10.1371/journal.pmed.1000217
17. Butcher NJ, Monsour A, Mew EJ, et al. Guidelines for reporting outcomes in trial protocols: the SPIRIT-Outcomes 2022 extension. *JAMA*. Published online December 13, 2022. doi:10.1001/jama.2022.21243
18. Butcher NJ, Mew EJ, Saeed L, et al. Guidance for reporting outcomes in clinical trials: scoping review protocol. *BMJ Open*. 2019;9(2):e023001.
19. 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
20. CONSORT. CONSORT statement website. Accessed March 31, 2022. <http://www.consort-statement.org/>
21. Ferreira-González I, Permyer-Miralda G, Busse JW, et al. Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. *J Clin Epidemiol*. 2007;60(7):651-657. doi:10.1016/j.jclinepi.2006.10.020
22. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ*. 2010;341:c3920. doi:10.1136/bmj.c3920
23. Cook JA, Hislop J, Adewuyi TE, et al. Assessing methods to specify the target difference for a randomised controlled trial: DELTA (Difference ELicitation in TriAls) review. *Health Technol Assess*. 2014;18(28). doi:10.3310/hta18280
24. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN checklist manual. Accessed February 28, 2022. [https://fac.ksu.edu/sites/default/files/cosmin\\_checklist\\_manual\\_v9.pdf](https://fac.ksu.edu/sites/default/files/cosmin_checklist_manual_v9.pdf)
25. Johnston BC, Ebrahim S, Carrasco-Labra A, et al. Minimally important difference estimates and methods: a protocol. *BMJ Open*. 2015;5(10):e007953. doi:10.1136/bmjopen-2015-007953
26. de Vet HC, Beckerman H, Terwee CB, Terluin B, Bouter LM. Definition of clinical differences. *J Rheumatol*. 2006;33(2):434-435.
27. de Vet HC, Terwee CB. The minimal detectable change should not replace the minimal important difference. *J Clin Epidemiol*. 2010;63(7):804-805. doi:10.1016/j.jclinepi.2009.12.015
28. Calvert M, Kyte D, Mercieca-Bebber R, et al; the SPIRIT-PRO Group. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO Extension. *JAMA*. 2018;319(5):483-494. doi:10.1001/jama.2017.21903
29. Hopewell S, Dutton S, Yu LM, Chan AW, 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
30. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745. doi:10.1016/j.jclinepi.2010.02.006
31. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine*. Cambridge University Press; 2011. doi:10.1017/CBO97805119996214
32. Mokkink L, Terwee C, de Vet H. Key concepts in clinical epidemiology: responsiveness, the longitudinal aspect of validity. *J Clin Epidemiol*. 2021;140:159-162. doi:10.1016/j.jclinepi.2021.06.002
33. Ferreira ML, Herbert RD, Ferreira PH, et al. A critical review of methods used to determine the smallest worthwhile effect of interventions for low back pain. *J Clin Epidemiol*. 2012;65(3):253-261. doi:10.1016/j.jclinepi.2011.06.018
34. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157. doi:10.1007/s11136-018-1798-3
35. Chan AW, 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
36. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med*. 2013;158(3):200-207. doi:10.7326/0003-4819-158-3-201302050-00583
37. Cook JA, Julious SA, Sones W, et al. Practical help for specifying the target difference in sample size calculations for RCTs: the DELTA<sup>2</sup> five-stage study, including a workshop. *Health Technol Assess*. 2019;23(60):1-88. doi:10.3310/hta23600
38. Cook JA, Julious 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-k3750. doi:10.1136/bmj.k3750
39. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database—update and key issues. *N Engl J Med*. 2011;364(9):852-860. doi:10.1056/NEJMsa1012065
40. OMERACT. OMERACT (Outcome Measures in Rheumatology): domain resources. Accessed May 19, 2022. <https://omeract.org/domain-selection/>
41. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med*. 2010;8(1):18. doi:10.1186/1741-7015-8-18
42. International Conference on Harmonisation. E3: structure and content of clinical study reports: step 4. Accessed February 28, 2022. <https://www.ich.org/page/efficacy-guidelines>
43. Williamson P, Altman D, Blazeby J, Clarke M, Gargon E. Driving up the quality and relevance of research through the use of agreed core outcomes. *J Health Serv Res Policy*. 2012;17(1):1-2. doi:10.1258/jhsrp.2011.011131
44. Page MJ, McKenzie JE, Forbes A. Many scenarios exist for selective inclusion and reporting of results in randomized trials and systematic reviews. *J Clin Epidemiol*. 2013;66(5):524-537. doi:10.1016/j.jclinepi.2012.10.010
45. Collister D, Bangdiwala S, Walsh M, et al. Patient reported outcome measures in clinical trials should be initially analyzed as continuous outcomes for statistical significance and responder analyses should be reserved as secondary analyses. *J Clin Epidemiol*. 2021;134:95-102. doi:10.1016/j.jclinepi.2021.01.026
46. Kravitz RL, Duan N; DEClDE Methods Center N-of-1 Guidance Panel, eds. *Design and Implementation of N-of-1 Trials: A User's Guide*. Agency for Healthcare Research and Quality; 2014.
47. Carrasco-Labra A, Devji T, Qasim A, et al. Minimal important difference estimates for patient-reported outcomes: a systematic survey. *J Clin Epidemiol*. 2021;133:61-71. doi:10.1016/j.jclinepi.2020.11.024
48. Woaye-Hune P, Hardouin JB, Lehur PA, Meurette G, Vanier A. Practical issues encountered while determining minimal clinically important difference in patient-reported outcomes. *Health Qual Life Outcomes*. 2020;18(1):156. doi:10.1186/s12955-020-01398-w
49. Moulhi Y, Jouve E, Castelli C, Gentile S. How is the minimal clinically important difference established in health-related quality of life instruments? review of anchors and methods. *Health Qual Life Outcomes*. 2020;18(1):136. doi:10.1186/s12955-020-01344-w
50. Devji T, Carrasco-Labra A, Guyatt G. Mind the methods of determining minimal important differences: three critical issues to consider. *Evid Based Ment Health*. 2021;24(2):77-81. doi:10.1136/ebmental-2020-300164
51. International Conference on Harmonisation. E9: statistical principles for clinical trials: current step 4 version. Accessed November 14, 2022. [https://database.ich.org/sites/default/files/E9\\_Guideline.pdf](https://database.ich.org/sites/default/files/E9_Guideline.pdf)

52. Mayo-Wilson E, Li T, Fusco N, et al. Cherry-picking by trialists and meta-analysts can drive conclusions about intervention efficacy. *J Clin Epidemiol*. 2017;91:95-110. doi:10.1016/j.jclinepi.2017.07.014
53. Butcher NJ, Mew E, Monsour A, et al. A scoping review of outcome-reporting recommendations for clinical trial protocols and reports. Accessed April 26, 2022. <https://osf.io/6f5x7/>
54. Eunetha (European Network for Health Technology Assessment). Guideline: endpoints used for relative effectiveness assessment of pharmaceuticals: composite endpoints. Accessed February 28, 2022. <https://www.eunetha.eu/wp-content/uploads/2018/01/Composite-endpoints.pdf>
55. Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Med*. 2015;13:282-282. doi:10.1186/s12916-015-0520-3
56. Campbell JM, Kulgar M, Ding S, et al. Diagnostic test accuracy systematic reviews. In: Aromataris EMZ, ed. *JBI Manual for Evidence Synthesis*. JBI; 2020. doi:10.46658/JBIMES-20-10
57. Mokkink LB, Prinsen CAC, Patrick DL, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs): user manual: version 1.0. Published 2018. Accessed November 14, 2022. [https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual\\_version-1\\_feb-2018.pdf](https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf)
58. De Los Reyes A, Augenstein TM, Wang M, et al. The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychol Bull*. 2015;141(4):858-900. doi:10.1037/a0038498
59. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. *J Natl Cancer Inst*. 2009;101(23):1624-1632. doi:10.1093/jnci/djp386
60. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource. Accessed May 19, 2022. <https://www.ncbi.nlm.nih.gov/books/NBK338448/>
61. Li G, Taljaard M, Van den Heuvel ER, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. *Int J Epidemiol*. 2017;46(2):746-755. doi:10.1093/ije/dyw320
62. European Medicines Agency. Guideline on multiplicity issues in clinical trials. Accessed February 8, 2022. [https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-multiplicity-issues-clinical-trials\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-multiplicity-issues-clinical-trials_en.pdf)
63. Gamble C, Krishan A, Stocken D, et al. Guidelines for the content of statistical analysis plans in clinical trials. *JAMA*. 2017;318(23):2337-2343. doi:10.1001/jama.2017.18556
64. Akl 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(12):e008431. doi:10.1136/bmjopen-2015-008431
65. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med*. 2012;367(14):1355-1360. doi:10.1056/NEJMs1203730
66. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions*, version 6.3. Accessed November 14, 2022. <http://www.training.cochrane.org/handbook>
67. Akl EA, Briel M, You JJ, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. *BMJ*. 2012;344:e2809. doi:10.1136/bmj.e2809
68. Rombach I, Rivero-Arias O, Gray AM, Jenkinson C, Burke Ó. The current practice of handling and reporting missing outcome data in eight widely used PROMs in RCT publications: a review of the current literature. *Qual Life Res*. 2016;25(7):1613-1623. doi:10.1007/s11136-015-1206-1
69. National Research Council (US) Panel on Handling Missing Data in Clinical Trials. *The Prevention and Treatment of Missing Data in Clinical Trials*. National Academies Press; 2010. doi:10.17226/12955
70. Little RJA. A test of missing completely at random for multivariate data with missing values. *J Am Stat Assoc*. 1988;83(404):1198-1202. doi:10.1080/01621459.1988.10478722
71. Kahan BC, Ahmad T, Forbes G, Cro S. Public availability and adherence to prespecified statistical analysis approaches was low in published randomized trials. *J Clin Epidemiol*. 2020;128:29-34. doi:10.1016/j.jclinepi.2020.07.015
72. Schriger DL. Guidelines for presenting tables and figures in scientific manuscripts. In: Moher D, Altman DG, Schulz KF, Simera I, Wager E, eds. *Guidelines for Reporting Health Research: A User's Manual*. Wiley; 2014. doi:10.1002/9781118715598.ch26
73. REDCap Research. CONSORT-Outcomes feedback form. Accessed February 28, 2022. <https://redcapexternal.research.sickkids.ca/surveys/?s=EL4LAR3HHR>
74. Qureshi R, Mayo-Wilson E, Li T. Harms in systematic reviews paper 1: an introduction to research on harms. *J Clin Epidemiol*. 2022;143:186-196. doi:10.1016/j.jclinepi.2021.10.023
75. Howick J, Webster R, Knottnerus JA, Moher D. Do overly complex reporting guidelines remove the focus from good clinical trials? *BMJ*. 2021;374:n1793. doi:10.1136/bmj.n1793