

Guidelines for Reporting Outcomes in Trial Reports The CONSORT-Outcomes 2022 Extension の和訳

著者 Nancy J. Butcher, PhD^{1,2}; Andrea Monsour, MPH¹; Emma J. Mew, MPH, MPhil^{1,3}; An-Wen Chan, MD, DPhil⁴; David Moher, PhD^{5,6}; Evan Mayo-Wilson, DPhil⁷; Caroline B. Terwee, PhD^{8,9}; Alyssandra Chee-A-Tow, MPH¹; Ami Baba, MRes¹; Frank Gavin, MA¹⁰; Jeremy M. Grimshaw, MBChB, PhD^{11,12}; Lauren E. Kelly, PhD^{13,14}; Leena Saeed, BSc, BEd¹; Lehana Thabane, PhD¹⁵; Lisa Askie, PhD¹⁶; Maureen Smith, MEd¹⁷; Mufiza Farid-Kapadia, MD, PhD¹; Paula R. Williamson, PhD¹⁸; Peter Szatmari, MD^{19,20}; Peter Tugwell, MD^{6,12,21,22}; Robert M. Golub, MD²³; Suneeta Monga, MD^{2,20}; Sunita Vohra, MD²⁴; Susan Marlin, MSc^{25,26}; Wendy J. Ungar, MSc, PhD^{1,27}; Martin Offringa, MD, PhD^{1,27,28}

翻訳 鈴木 直子 (SUZUKI Naoko)^{1*}, 野田 和彦 (NODA Kazuhiko)¹, 波多野 絵梨 (HATANO Eri)¹, 高橋 德行 (TAKAHASHI Noriyuki)¹, 中村 駿一 (NAKAMURA Shunichi)¹, LIU XUN¹, 柿沼 俊光 (KAKINUMA Toshihiro)¹, 馬場 亜沙美 (BABA Asami)¹, 山本 和雄 (YAMAMOTO Kazuo)¹

Key Words: ヒト臨床試験, ヒト試験, 特定保健用食品, 機能性表示食品, 研究デザイン, CONSORT, ガイドライン, アウトカム, SPIRIT

所属 (著者)

¹ Child Health Evaluative Sciences, The Hospital for Sick Children Research Institute, Toronto, Ontario, Canada

² Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

³ Department of Chronic Disease Epidemiology, School of Public Health, Yale University, New Haven, Connecticut

⁴ Department of Medicine, Women's College Research Institute, University of Toronto, Toronto, Ontario, Canada

⁵ Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

⁶ School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada

⁷ Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill

⁸ Amsterdam University Medical Centers, Vrije Universiteit, Department of Epidemiology and Data Science, Amsterdam, the Netherlands

⁹ Department of Methodology, Amsterdam Public Health Research Institute, Amsterdam, the Netherlands

¹⁰ public panel member, Toronto, Ontario, Canada

¹¹ Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

¹² Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada

¹³ Department of Pharmacology and Therapeutics, University of Manitoba, Winnipeg, Canada

¹⁴ Children's Hospital Research Institute of Manitoba, Winnipeg, Canada

¹⁵ Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

¹⁶ NHMRC Clinical Trials Centre, University of Sydney, Sydney, New South Wales, Australia

¹⁷ patient panel member, Ottawa, Ontario, Canada

¹⁸ MRC-NIHR Trials Methodology Research Partnership, Department of Health Data Science, University of Liverpool, Liverpool, England

¹⁹ Cundill Centre for Child and Youth Depression, Centre for Addiction and Mental Health, Toronto, Ontario, Canada

²⁰ Department of Psychiatry, The Hospital for Sick Children, Toronto, Ontario, Canada

²¹ Bruyère Research Institute, Ottawa, Ontario, Canada

²² Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

²³ Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois

²⁴ Departments of Pediatrics and Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Canada

²⁵ Clinical Trials Ontario, Toronto, Canada

²⁶ Department of Public Health Sciences, Queen's University, Kingston, Ontario, Canada

²⁷ Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, Ontario, Canada

²⁸ Division of Neonatology, The Hospital for Sick Children, Toronto, Ontario, Canada

* 責任著者: 鈴木 直子

所属 (翻訳)

¹ 株式会社オルトメディコ

〒112-0002 東京都文京区小石川 1-4-1 住友不動産後楽園ビル 2 階

Japanese translation of “Guidelines for Reporting Outcomes in Trial Reports: The CONSORT-Outcomes 2022 Extension”

Authors

Nancy J. Butcher, PhD^{1,2}; Andrea Monsour, MPH¹; Emma J. Mew, MPH, MPhil^{1,3}; An-Wen Chan, MD, DPhil⁴; David Moher, PhD^{5,6}; Evan Mayo-Wilson, DPhil⁷; Caroline B. Terwee, PhD^{8,9}; Alyssandra Chee-A-Tow, MPH¹; Ami Baba, MRes¹; Frank Gavin, MA¹⁰; Jeremy M. Grimshaw, MBChB, PhD^{11,12}; Lauren E. Kelly, PhD^{13,14}; Leena Saeed, BSc, BEd¹; Lehana Thabane, PhD¹⁵; Lisa Askie, PhD¹⁶; Maureen Smith, MEd¹⁷; Mufiza Farid-Kapadia, MD, PhD¹; Paula R. Williamson, PhD¹⁸; Peter Szatmari, MD^{19,20}; Peter Tugwell, MD^{6,12,21,22}; Robert M. Golub, MD²³; Suneeta Monga, MD^{2,20}; Sunita Vohra, MD²⁴; Susan Marlin, MSc^{25,26}; Wendy J. Ungar, MSc, PhD^{1,27}; Martin Offringa, MD, PhD^{1,27,28}

Translator

Naoko Suzuki^{1*}, Kazuhiko Noda¹, Eri Hatano¹, Noriyuki Takahashi¹, Shunichi Nakamura¹, Xun Liu¹, Toshihiro Kakinuma¹, Asami Baba¹, Kazuo Yamamoto¹

*Correspondence author: Naoko Suzuki

Keywords: Human clinical trial, Human trials, Food for specified health uses, Food with function claims, Research Design, CONSORT, guideline, outcome, SPIRIT

Affiliation (Author)

¹ Child Health Evaluative Sciences, The Hospital for Sick Children Research Institute, Toronto, Ontario, Canada

² Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

³ Department of Chronic Disease Epidemiology, School of Public Health, Yale University, New Haven, Connecticut

⁴ Department of Medicine, Women’s College Research Institute, University of Toronto, Toronto, Ontario, Canada

⁵ Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

⁶ School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada

⁷ Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill

⁸ Amsterdam University Medical Centers, Vrije Universiteit, Department of Epidemiology and Data Science, Amsterdam, the Netherlands

⁹ Department of Methodology, Amsterdam Public Health Research Institute, Amsterdam, the Netherlands

¹⁰ public panel member, Toronto, Ontario, Canada

¹¹ Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

¹² Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada

¹³ Department of Pharmacology and Therapeutics, University of Manitoba, Winnipeg, Canada

¹⁴ Children’s Hospital Research Institute of Manitoba, Winnipeg, Canada

¹⁵ Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

¹⁶ NHMRC Clinical Trials Centre, University of Sydney, Sydney, New South Wales, Australia

¹⁷ patient panel member, Ottawa, Ontario, Canada

¹⁸ MRC-NIHR Trials Methodology Research Partnership, Department of Health Data Science, University of Liverpool, Liverpool, England

¹⁹ Cundill Centre for Child and Youth Depression, Centre for Addiction and Mental Health, Toronto, Ontario, Canada

²⁰ Department of Psychiatry, The Hospital for Sick Children, Toronto, Ontario, Canada

²¹ Bruyère Research Institute, Ottawa, Ontario, Canada

²² Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

²³ Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois

²⁴ Departments of Pediatrics and Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Canada

²⁵ Clinical Trials Ontario, Toronto, Canada

²⁶ Department of Public Health Sciences, Queen’s University, Kingston, Ontario, Canada

²⁷ Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, Ontario, Canada

²⁸ Division of Neonatology, The Hospital for Sick Children, Toronto, Ontario, Canada

Affiliated institution

¹ ORTHOMEDICO Inc.

2F Sumitomo Fudosan Korakuen Bldg., 1-4-1 Koishikawa, Bunkyo-ku, Tokyo, 112-0002, Japan.

本項について

本稿は、EQUATOR Network が提供するガイドラインの一つである「Guidelines for Reporting Outcomes in Trial Reports The CONSORT-Outcomes 2022 Extension」の本文を翻訳したものである。なお、付録は原文（doi: 10.1001/jama.2022.21022）にアクセスして参照する。また、表 1 は「Guidelines for Reporting Outcomes in Trial Reports The CONSORT-Outcomes 2022 Extension」で拡張された項のみを抽出した。

キーポイント

問題

出版された臨床試験報告書には、どのようなアウトカム特有の情報を含めるべきか？

結果

Enhancing the Quality and Transparency of Health Research (EQUATOR) の方法論的枠組みの手法を応用したエビデンスおよび国際的なコンセンサスに基づく方法により、17 のアウトカム特有の報告項目が特定された。

意義

臨床試験報告書にこれらの項目を含めることは、試験の有用性、再現性、透明性を高め、試験結果の選択的不報告を制限するのに役立つ可能性がある。

抄録

重要性

臨床医、患者、政策立案者は、エビデンスに基づいた意思決定を行うために、出版された臨床試験結果に頼っている。臨床試験結果を批判的に評価し利用するためには、読者は、何が計画され、実施され、発見されたかに関する完全で透明性のある情報を必要とする。アウトカムの選択、評価、解析の問題をあいまいにするような不十分な報告方法を減らすためには、臨床試験の出版物においてどのようなアウトカム特有の情報を報告すべきかについて、具体的かつ統一されたガイダンスが必要である。

目的

Consolidated Standards of Reporting Trials (CONSORT) 2010 声明との統合により、臨床試験報告書におけるアウトカムの報告について、エビデンスおよびコンセンサスに基づいた、かつ統一された基準を作成する。

エビデンスレビュー

Enhancing the Quality and Transparency of Health Research (EQUATOR) の方法論的枠組みを用いて、CONSORT 2010 声明の CONSORT-Outcomes 2022 拡張版は、下記の通り、作成された。

- (1) 専門家との協議および臨床試験のアウトカム報告に関する既存のガイダンス（2018年3月19日以前10年以内に公表されたもの）のスコopingレビューにより、アウトカム報告の候補項目を作成・評価した。既存のガイダンスは、専門家からの意見収集、MEDLINE と Cochrane Methodology Register での電子データベース検索、灰色文献検索、参考文献リスト検索により特定された。
- (2) 22 ヶ国 124 名のパネリストにより、国際的なデルファイ調査（2018年11月～2019年2月）が3回実施され、追加項目を評価・特定した。
- (3) 25 名のパネリストが参加した対面コンセンサス会議（2019年4月9日～10日）を開催し、臨床試験報告書におけるアウトカムの報告の必須項目を特定した。

結果

スコーピングレビューおよび専門家との協議により、臨床試験報告書におけるアウトカムの報告に関連する128の推奨事項が特定され、その大部分(83%)はCONSORT 2010声明に含まれていなかった。すべての推奨事項は、デルファイ調査を行うために64項目に統合整理され、デルファイ調査の結果、30項目がコンセンサス会議でさらに評価し、CONSORT-Outcomes 2022 拡張版に含める可能性がある基準を満たした。コンセンサス会議の期間中および終了後の議論により、CONSORT 2010 声明チェックリストの項目をより詳細にし、下記の内容に関連する17項目が得られた。いつ、どのように評価したかを含め、試験アウトカムの完全な定義付けと妥当性確認(CONSORT 2010 声明チェックリスト項目6a)、サンプルサイズ算出時の治療群間における目標差(target difference)の定義付けと妥当性確認(CONSORT 2010 声明チェックリスト項目7a)、主要・副次的アウトカムの群間比較に用いた統計的手法の記述(CONSORT 2010 声明チェックリスト項目12a)、事前に決められた解析と事前に決められなかったアウトカム解析の記述(CONSORT 2010 声明チェックリスト項目18)。

結論と関連性

本CONSORT 2010 声明のCONSORT-Outcomes 2022 拡張版は、すべての出版された臨床試験報告書で検討すべき17のアウトカム特有の報告項目を提供し、試験の有用性、再現性、透明性の向上に役立ち、試験結果の選択的不報告のリスクを最小限に抑える可能性がある。

はじめに

十分にデザインされ、適切に実施されたランダム化臨床試験(RCT)は、エビデンスに基づく臨床的意思決定に役立つ一次エビデンスを得るためのゴールドスタンダードである。RCTでは、参加者に対する介入効果を評価するために試験アウトカムが使用される¹⁾。Consolidated Standards of Reporting Trials (CONSORT) 2010 声明では、出版されたRCT報告に含めるべき25の報告項目が示されている^{2,3)}。

試験アウトカムの完全な報告は、結果の再現性、知識の合成(knowledge synthesis)、結果の選択的不報告を防ぐために重要である。スコーピングレビューの結果、アカデミア、規制当局、その他の主要な情報源により、出版された報告書における試験アウトカムの報告方法に関する多様で一貫性のない推奨事項が明らかになった⁴⁾。不十分なアウトカム報告は、様々な学術ジャーナルや専門分野でよく見られており、アウトカムの選択、定義、評価、分析、および事前に決められたアウトカム(すなわち、プロトコルや試験登録の内容)からの変更に関する重要な情報は、しばしば十分に報告されていない⁵⁻⁹⁾。このような回避可能な報告の問題は、システマティックレビューやメタアナリシスから導かれ

る結論に影響を与えることが示されており¹⁰⁾、研究を無駄にする一因となっている¹¹⁾。

試験アウトカムの報告に改善が求められているが^{5,12)}、試験担当者、システマティックレビュー作成者、ジャーナル編集者、臨床医、患者、一般市民などの知識利用者にとって、何が有用で完全な試験アウトカム報告であるかは明確になっていない⁴⁾。特定の種類の試験アウトカムの報告に関連する2つのCONSORT 声明の拡張版(2004年に「害」、2013年に「患者報告アウトカム」)が存在している^{13,14)}。しかし、すべてのアウトカムの種類、集団、試験デザインに適用可能なアウトカム特有の必須情報に関する標準的な報告ガイドラインは存在しない⁴⁾。

CONSORT-Outcomes 2022 拡張版の目的は、臨床試験報告のために、統一された、エビデンスおよびコンセンサスに基づいたアウトカム報告基準を作成することであった。

方法

CONSORT-Outcomes 2022 拡張版は、Instrument for Reporting Planned Endpoints in Clinical Trials (InsPECT) プロジェクト¹⁵⁾の一環として、

Enhancing the Quality and Transparency of Health Research (EQUATOR) の報告ガイドライン作成のための方法的枠組みに従って作成された¹⁶⁾。The Hospital for Sick Children の研究倫理委員会の決定により、倫理承認は必要とされなかった。CONSORT-Outcomes 2022 拡張版の作成¹⁵⁾は、臨床試験プロトコルに関する Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)-Outcomes 2022 拡張版と並行して行われた¹⁷⁾。

主な作成段階

第1に、専門家との協議および試験アウトカムの報告に関する既存のガイダンス（2018年3月19日以前10年以内に公表されたもの）のスコوپングレビューの結果を統合し、RCTのアウトカム報告に関連する推奨事項の初期リストを作成した^{4,18)}。既存のガイダンスは、専門家からの意見収集、MEDLINEとCochrane Methodology Registerでの電子データベース検索、灰色文献検索、参考文献リスト検索により特定された。第2に、2018年11月から2019年2月にかけて、追加項目の特定および各項目の重要性評価のために、国際的なデルファイ調査が3回実施された。この調査は、22カ国124人のパネリストによって行われた（付録のeTable 1）。第3に、患者パートナーと公的パートナーを含む4カ国25人のパネリストが対面で参加した専門家コンセンサス会議が開催され（2019年4月9日～10日）、試験報告のアウトカム報告に関連する必須項目を特定し、普及活動を確立した。項目の選択と文言は、幹部パネリストによるコンセンサス後会議および、コンセンサス会議のパネリストによる電子メール確認を通じて最終決定された。

その他の情報

CONSORT-Outcomes 2022 拡張版の作成に関する詳細な方法は、各段階で評価された項目数および最終項目決定までのプロセス（付録のeFigure）を含め、付録のeAppendix 1に記載されている。スコوپングレビューのプロトコルと結果はすでに出版されており^{4,18)}、付録のeAppendix 1に示されている。検索戦略は付録のeAppendix 2に記載されている。デルファイ調査のパネリストおよびコンセンサス会議のパネリストの自己申告による特徴は、付録

のeTable 1-2に記載されている。患者パートナーおよび公的パートナーの関与に関する詳細は、付録のeAppendix 1に記載されている。

結果

CONSORT 2010 声明のチェックリスト項目に加え、CONSORT-Outcomes 2022 拡張版では、試験デザインや集団に関係なく、臨床試験報告書に最低17のアウトカム特有の報告項目を含めることが推奨されている。スコوپングレビューおよび専門家との協議により、臨床試験報告書におけるアウトカムの報告に関連する128の推奨事項が特定されたが、その大部分（83%）はCONSORT 2010 声明に含まれていなかった。すべての推奨事項は、デルファイ法調査を行うために64項目に統合整理され、デルファイ調査の結果、30項目がコンセンサス会議でさらに評価し、CONSORT-Outcomes 2022 拡張版に含める可能性がある基準を満たした。CONSORT 2010 声明チェックリスト項目およびCONSORT-Outcomes 2022 拡張版で追加された17項目は、表1に示されている¹⁹⁾

チェックリストの記入可能バージョンは、付録のeTables 3-4 およびCONSORT ウェブサイトに掲載されている²⁰⁾。更新されたチェックリストを使用するには、専門分野や地域によって、用語や定義が異なる場合があるため、用語集²¹⁻³⁸⁾ (Box) の主要用語の定義を参照する必要がある。定義されたアウトカムの5つのコア要素（例付き）は、表2に示されている^{39,40)}。

CONSORT-Outcomes 2022 拡張版の新しいチェックリスト項目をCONSORT 2010 声明と合わせて適用することにより、臨床試験のアウトカムが完全に定義され、かつ報告されることが担保される。17のCONSORT-Outcomes 2022 拡張版のチェックリスト項目を実施することで、主要なユーザー、ユーザーの行動案、および結果として生じた潜在的な利点の推定リストは、付録のeTable 5に記載されており、コンセンサス会議のナレッジ・トランスレーション(knowledge translation) セッションより作成された。これらのアウトカム報告に関する推奨事項の検討と適用は、試験論文の著者、ジャーナル編集者、査読者、システムティックレビュー作成者、患者、一般

表 1 CONSORT2010 声明および CONSORT-Outcomes 2022 拡張版における臨床試験報告の推奨チェックリスト項目^{a,b} (1/2)

章 (Section)	項目番号 (Item No.)	CONSORT 2010 声明	項目番号 (Item No.)	CONSORT-Outcomes 2022 拡張版
方法 [Methods]				
アウトカム [Outcomes]	6a	事前に特定され明確に定義された主要・副次的アウトカム評価項目。いつどのように評価されたかを含む。	6a.1	試験の主要アウトカムのドメインを選択した根拠を示す。
			6a.2	特定の測定変量 (例: 収縮期血圧), 解析指標 (例: ベースラインからの変化, 最終値, 事象発生までの時間), 集計方法 (例: 平均値, 割合), および各アウトカムの評価時点を記述する。
			6a.3	主要アウトカムの解析指標が参加者内の変化を表す場合, 個人における最小重要変化量を定義し, 正当化する。
			6a.4	アウトカムのデータが連続的であるが, カテゴリカルデータとして解析された場合 (集計方法), 使用されたカットオフ値を明記する。
			6a.5	ランダム化後, 複数の時点でアウトカム評価が実施された場合は, 解析に使用された評価時点を記載する。
			6a.6	複合アウトカムを使用した場合は, 複合アウトカムの個別構成要素をすべて定義する。
			6a.7	試験登録または試験プロトコルで事前に決められていないアウトカムを特定する。
			6a.8	アウトカムを評価するために使用された評価尺度 (例: 質問票, 臨床検査) を, それらの試験参加者と同様の集団における信頼性, 妥当性, 応答性とともに記述する。
			6a.9	アウトカムを評価した人 (例: 看護師, 保護者) と, アウトカムを評価するために評価尺度の実施に必要な資格または試験に特化した訓練を記述する。
			6a.10	データ収集中 (例: 重複測定) およびデータ収集後 (例: アウトカムデータの値の範囲検査) にアウトカムデータの質を高めるために使用されたプロセスを記述する, またはそれらの詳細を確認できる場所を記載する。
	6b	試験開始後のアウトカムの変更とその理由。		
症例数 [Sample size]	7a	どのように目標症例数が決められたか。	7a.1	治療群間の目標差 (例: 最小重要差) を定義し, 正当化する。
	7b	あてはまる場合には, 中間解析と中止基準の説明。		
統計学的手法	12a	主要・副次的アウトカムの群間比較に用いられた統計学的手法。	12a.1	主要および副次的アウトカムの解析または解釈における多重性を説明するために使用した手法を記述する (例: 共主要アウトカム, 複数の時点で評価された同一アウトカム, アウトカムのサブグループ解析)。
			12a.2	解析および報告からアウトカムデータを除外する基準を記載し, 正当化する, またはアウトカムデータを除外しなかったことを報告する。

表 1 CONSORT2010 声明および CONSORT-Outcomes 2022 拡張版における臨床試験報告の推奨チェックリスト項目^{a,b} (2/2)

章 (Section)	項目番号 (Item No.)	CONSORT 2010 声明	項目番号 (Item No.)	CONSORT-Outcomes 2022 拡張版
			12a.3	欠測のパターン (例: ランダムでない欠測) を評価するために使用された方法, およびアウトカム項目または評価全体の欠測を対処するために使用された方法を記述する。
			12a.4	試験プロトコルの不遵守に関連するアウトカム解析集団の定義を示す (例: ランダム化解析として)。
	12b	サブグループ解析や調整解析のような追加的解析の手法。		
結果 [Results]				
アウトカムと推定 [Outcomes and Estimation]	17a	主要・副次的アウトカムのそれぞれについて, 各群の結果, 介入のエフェクト・サイズの推定とその精度 (95% 信頼区間など)。	17a.1	事前に決められたすべてのアウトカム解析の結果を含める, または本報告書に記載がない場合はどこに結果の記載があるかを示す。
	17b	2 項アウトカムについては, 絶対エフェクト・サイズと相対エフェクト・サイズの両方を記載することが推奨される。		
補助的分析 [Ancillary analyses]	18	サブグループ解析や調整解析を含む, 実施した他の解析の結果。事前に特定された解析と探索的解析を区別する。	18.1	事前に決められていない解析があった場合は, その理由を説明する。

略語: CONSORT = Consolidated Standards of Reporting Trials

a: このチェックリストは, チェックリスト項目に関する重要な説明を行うために, CONSORT 2010 声明のガイドライン^{2,3)}と合わせて確認することが強く推奨されている。CONSORT 2010 声明のチェックリストは, クリエイティブ・コモンズ・ライセンスの条件の下で配布されている。

b: 「Guidelines for Reporting Outcomes in Trial Reports The CONSORT-Outcomes 2022 Extension」で拡張された項目のみを抽出した。

市民, 試験参加者にとって役立つだろう (付録の eTable 5)。

本報告書には, CONSORT-Outcomes 2022 拡張版から作られた 17 のチェックリスト項目の簡単な説明が含まれている。既存のチェックリスト項目の報告方法に関するガイダンスは, 表 1 (本稿では, 拡張された項のみを抽出しているため, 変更のない項に関しては, 原文の Table 1 を参照する) の CONSORT 2010 声明²⁾ およびガイドライン解説書⁴¹⁾に記載されている。一部の試験報告書や試験関連書類 (統計解析計画書や臨床研究レポート⁴²⁾など) に含めると役立つだろう追加項目は, 付録の eTable 6 に示されているが, すべての試験報告書に必須の報告項目とは考えられていない。

臨床試験で使用されるアウトカムの選択, 測定, および記述の方法に関する CONSORT-Outcomes 2022 拡張版のチェックリスト項目

項目 6a.1. 試験の主要アウトカムのドメインを選択した根拠を示す。

この項目は, CONSORT 2010 声明のチェックリスト項目 6a を拡張し, 主要アウトカムとして使用するアウトカムドメインの選択の根拠に関する報告を明確に求めるものである。大まかな概念レベルでは, アウトカムのドメインは, アウトカムを記述するために使用される名称や概念 (例: 疼痛) を指している^{10,39)}。「ドメイン」という用語は, 文献では「コンストラクト (construct)」や「アトリビュート (attribute)」といった用語と密接に関連し, 同等に用いられることもある⁴⁰⁾。完全なアウトカムの定義は試験報告書に示されることが期待されている

表 2 定義されたアウトカムの5つのコア要素^a

番号 [Element No.]	用語 [Element term]	使用定義 [Definition used]	例 1 [Example 1]	例 2 [Example 2]	例 3 [Example 3]
1	ドメイン ^b	1つ以上のアウトカムを説明するタイトルまたは概念	血圧	うつ病	死亡
2	測定変量または特定の測定	試験参加者から直接収集されたデータと一致する；アウトカムドメインを評価するために使用された評価尺度を含む説明。			
		説明的名称	オムロンの上腕式血圧計で測定された収縮期血圧	MADRS	病院データベースごとの全死因死亡率
		該当する場合、解析しようとする合計スコアまたは下位尺度	該当しない	MADRS 合計スコア	該当しない
3	特定の解析指標	解析における参加者レベルの測定単位（例：ベースラインからの変化量、最終値またはある評価時点における値、事象発生までの時間 (time-to-event)）。	ある時点の値	ベースラインからの変化量	事象発生までの時間
4	集計方法	治療効果の推定手順			
		● アウトカムを連続変数、カテゴリカル変数、time-to-event型変数とする場合	連続変数	二値変数	事象発生までの時間
		● 連続変数の場合、中心傾向の測定値（例：平均値）。カテゴリカル変数およびtime-to-event型のデータの場合、事象を有する割合と、該当する場合には、比較された特定のカットオフ値またはカテゴリー。	平均値	50%以上の減少を示した参加者の割合	罹患密度 (incidence density) および群間の罹患密度率 (incidence density rate)
5	評価時点	追跡調査の測定時期			
		● いつアウトカム測定を行うか	ランダム化から 2, 4, 12 週後	ランダム化から 2, 4, 6, 8 週後	毎日
		● どのアウトカム測定を解析するか	ランダム化から 12 週後	ランダム化から 8 週後	追跡終了

略語：MADRS = Montgomery-Åsberg Depression Rating Scale

a: 内容は, Zarin *et al.*³⁹⁾, Mayo-Wilson *et al.*¹⁰⁾, Chan *et al.*³⁵⁾ を参考にした。

b: 試験のアウトカムを定義する際には、必要に応じて、試験プロトコルにアウトカムドメインの明示的かつ具体的な説明を示すべきである。疼痛などアウトカムドメインの幅が広い場合、プロトコルで定められた特定のドメインは、ある通常の日の24時間において、無痛から想像しうる最悪の痛みまでの範囲で表される痛みの強さの日間平均値に定義されることが考えられる⁴⁰⁾。

が (CONSORT 2010 声明チェックリスト項目 6a で推奨されているように)^{2, 40)}、試験の主要アウトカムのアウトカムドメインを選択した根拠も、計画された試験の目的を裏付けるものであるため、報告す

ることが不可欠である。

根拠となる重要な点には、(1)試験に関与する人々 (例：患者、一般市民、臨床医、政策立案者、資金

提供者、医療費負担者) にとってのアウトカムドメインの重要性、(2) アウトカムドメインに対する介入の期待効果、(3) 試験中にアウトカムドメインを正確、安全、かつ実行可能な方法で評価する能力が含まれる。また、選択されたアウトカムドメインがコアアウトカムセット (core outcome set) (すなわち、特定の臨床領域において、合意されたすべての試験で測定されるべき標準アウトカムセット) に由来するかどうかも報告されるべきである⁴³⁾。

項目 6a.2. 特定の測定変量 (例: 収縮期血圧)、解析指標 (analysis metric) (例: ベースラインからの変化量、最終値、事象発生までの時間)、集計方法 (例: 平均値、割合)、および各アウトカムの評価時点を記述する

この項目は、事前に決められた主要および副次的アウトカムを完全に定義することを推奨する CONSORT 2010 声明のチェックリスト項目 6a を拡張したものであり、SPIRIT 2013 声明のチェックリスト項目 12 (および試験アウトカムの定義に関する説明文) を反映した具体的な推奨事項を示している³⁵⁾。CONSORT-outcomes 2022 拡張版のチェックリスト項目 6a.2 では、測定変量、特定の解析指標、集計方法、評価時点など、アウトカムの各要素を記述することが推奨されている。ClinicalTrials.gov のような試験登録機関では、すでにこの枠組みを用いてアウトカムを定義することが求められている^{10,35,39)}。試験アウトカムの各要素を明確に報告しないことは、検知不能な多重検定、データのチェリー・ピッキング (cherry picking)、計画とは異なる試験報告書における結果の選択的不報告を可能にする^{10,44)}。

項目 6a.3. 主要アウトカムの解析指標が参加者内の変化を表す場合、個人における最小重要変化量を定義し、正当化する。

この項目は、CONSORT-Outcomes 2022 拡張版のチェックリスト項目 6a.2. を拡張したものである。主要アウトカムの参加者レベルの解析指標が、以前の値 (ベースライン時の測定値など) からの個人内変化を表す場合、関連する評価尺度の最小重要変化量 (MIC) と考えられるものの定義と正当性を示すべきである。CONSORT-Outcomes2022 拡張版では、MIC は患者、臨床医、または試験関係者が重要と

考える最小の患者内変化量と定義された (一般的な代替用語は Box に示されている)^{24,25,31)}。MIC は、各治療群で MIC 値以上の変化を示した参加者の割合 (例えば、改善した割合を定義するため)⁴⁵⁾ を解析する予定の試験や n-of-1 試験など、参加者内変化の指標を用いるすべての試験で報告されることが重要である⁴⁶⁾。

Box. CONSORT- Outcomes 2022 拡張版で使用された用語集

複合アウトカム: 複合アウトカムは、2つ以上の構成アウトカム (例: 死亡または非致命的脳卒中を経験した参加者の割合) から構成される。構成アウトカムで指定された事象のいずれかを体験した参加者は、複合アウトカムを経験したとみなされる^{21,22)}。

CONSORT 2010: 2010 年に発表された Consolidated Standards of Reporting Trials (CONSORT) 声明^{2,3)}。

CONSORT-Outcomes 2022 拡張版: CONSORT 2010 声明ではカバーされていない、アウトカムに関連する内容を記述した追加必須チェックリスト項目。

構成概念妥当性: 評価尺度が測定対象ドメインを妥当に測定しているという仮定に基づき、試験で報告された得点が仮説 (例: 内部関係、得点と他の尺度の関係、関連する集団間の差異に関するもの) と一致する度合い³⁰⁾。

内容妥当性: 評価尺度の内容が測定対象ドメインを適切に反映している度合い³⁰⁾。

基準妥当性: ある評価尺度の得点がゴールドスタンダードを適切に反映している度合い³⁰⁾。

異文化間の妥当性: 翻訳された、または文化的適応された評価尺度の項目の性能がその評価尺度の原版を使用した場合の当該項目の性能を適切に反映している度合い³⁰⁾。

最小重要変化量: 患者、臨床医、その他関係者が重要であると考え、患者内の最小の変化量^{4,5)}。その変化量は、スコアや測定単位 (連続の測定値や

順序型の測定値), または頻度 (二値アウトカム) の変化量である。この用語は, 健康保健分野の文献では, しばしば最小重要差 (minimal important difference) という用語と互換的に使用されている。CONSORT-Outcomes 2022 拡張版では, 概念上, 「最小重要変化量」は重要な患者内変化量 (項目 6a.3) を指している一方, 「最小重要差」は重要な群間差を指している。「minimal」の代わりに「minimum」を用いたり, 「clinically」や「clinical」などの形容詞を付けたりするなどの細かい用語の変形はよく見られる (例: the minimum clinically important change) ²³⁾。

最小重要差: 患者, 臨床医, その他関係者が重要であると考え, 最小の群間差 ^{24,27)}。その差は, スコアや測定単位 (連続の測定値や順序型の測定値) または頻度 (二値アウトカム) の差である。「minimal」の代わりに「minimum」を用いたり, 「clinically」や「clinical」などの形容詞を付けたりするなどの細かい用語の変形はよく見られる (例: the minimum clinically important difference) ²³⁾。

アウトカム: 健康介入への曝露の効果を検討するために評価されるものを指す ¹⁾。定義されたアウトカムの5つのコア要素は, 表2に示されている。

主要アウトカム: 試験の主要目的に最も直接的に関連する計画されたアウトカム ²⁸⁾。これは通常, 効力 (efficacy) や有効性 (effectiveness) の評価を主要目的とする試験のサンプルサイズ計算に使用される結果である ²⁹⁾。多くの試験は主要アウトカムが1つであるが, 1つ以上ある試験もある。医学文献では, 主要アウトカムを指す場合, 「primary end point」という用語が使われることがある ⁴⁾。

信頼性: 測定に誤差がない度合い。具体的には, いくつかの条件下で (例: 内的整合性を調べるために同じ評価尺度の異なる項目セットを使用すること, 経時的測定 (over time) や再検査法 (test-retest), 同じ場面や評価者間 (interrater) で異なる人物によるもの, または異なる場面や評価者内 (intrarater) で, 評価者や回答者などの同じ人物によるもの), 参加者の得点に変化せず, 繰り返し測定した場合に同じ

である程度 ³⁰⁾。

応答性: アウトカム領域における経時的变化を正確に検出し測定するための評価尺度の能力 ^{31,32)}。単一スコアの妥当性を指す評価尺度の構成概念妥当性や基準妥当性とは異なり, 応答性は変化スコアの妥当性 (すなわち縦断的妥当性) を指す ³⁰⁾。

副次的アウトカム: 介入の付加的な効果を評価するために, 試験プロトコルで事前に決められたアウトカム ²⁸⁾。

最小有益効果: 患者によって決められた介入のコスト, 潜在的な害, 不便さを正当化できる, 介入による最小の有益な効果 ³³⁾。

SPIRIT 2013: 2013年に発表された Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 声明 ^{35,36)}。

SPIRIT-Outcomes 2022 拡張版: SPIRIT 2013 声明ではカバーされていない, アウトカムに関連する試験プロトコルの内容を説明した追加必須チェックリスト項目 ¹⁷⁾。

構造的妥当性: 評価尺度 (例: 患者質問票や臨床評価尺度) の得点が, 測定されるドメインの次元性を適切に反映している度合い ³⁰⁾。

評価尺度: 評価を行うために使用される尺度またはツール。評価尺度は, 質問票, 臨床評価尺度, 臨床検査, 身体検査や画像観察によって得られるスコア, または1つの質問に対する回答である可能性がある ³⁴⁾。

目標差: サンプルサイズの計算において, 介入群間の主要アウトカムで検出される差として使用され, 1つ以上の主要な利害関係者グループによって現実的または重要である (最小重要差や最小有益効果など) とみなされる値 ^{37,38)}。

妥当性: ある評価尺度が, 測定しようとしているドメインを測定する度合い ³⁰⁾。

MICを記述することで、患者、臨床医、政策立案者が試験結果とその臨床的関連性を理解しやすくなる。臨床試験の知識があるユーザーは、ベンチマークとしてのMIC自体に興味を持つかもしれないし、代わりに、既知のMICよりも大きな値に興味を持つかもしれない。選択されたMICの正当性を記述することは重要である。その理由は、同じ評価尺度に対して複数のMICが存在する可能性があり、その臨床的関連性や方法論的質は、どのように、誰を対象として決定されたかによって異なるからである⁴⁷⁻⁵⁰。試験の母集団と設定に対してMICが不明な場合は、その旨を報告すべきである。

項目 6a.4. アウトカムのデータが連続的であるが、カテゴリカルデータとして解析された場合（集計方法）、使用されたカットオフ値を明記する。

この項目は、CONSORT-Outcomes 2022 拡張版のチェックリスト項目 6a.2 を拡張したもので、事前に決められ、連続尺度（または順序尺度）で収集されたアウトカムのデータを解析のためにカテゴリカル変数に変換する際に使用されたカットオフ値について説明するよう、必要に応じて著者に促すものである^{10,35}。カットオフ値の選択の根拠を説明することが推奨される。異なる試験で異なるカットオフ値が適用されることは稀ではないのである。選択されたカットオフ値が最も有用なのは、臨床的関連性が明確な場合である⁵¹。この情報を報告することで、多重検定（「p 値ハッキング（p-hacking）」として知られる）、データのチェリー・ピッキング、試験報告書における結果の選択的不報告を避けることができる^{10,44,52}。

項目 6a.5. ランダム化後、複数の時点でアウトカム評価が実施された場合は、解析に使用された評価時点を記載する。

この項目は、CONSORT-Outcomes 2022 拡張版のチェックリスト項目 6a.2 の評価時点に関する内容を拡張したもので、アウトカム評価がランダム化後に複数の時点で実施された場合（例：ランダム化後の12週間の間に、毎日血圧を評価した試験）、主解析で使用された評価時点を明記するよう、必要に応じて著者に促すものである。解析に使用する評価時点を事前に計画した上で明記することで、予定外の

複数の評価時点における解析、大きな結果や有意な結果が得られなかった評価時点の選択的不報告の可能性を抑えることができる^{35,39}。

評価時点の選択に根拠を示すことが奨励される（例：予想される介入後の臨床経過（clinical trajectory）や、臨床的に意味のある治療効果を達成するために必要な治療期間に基づく）。追跡期間の長さは、試験で検討する予定の治療方針（management decision）に対して適切でなければならない⁵³。

項目 6a.6. 複合アウトカムを使用した場合は、複合アウトカムの個別構成要素をすべて定義する。

複合アウトカムは、関連する可能性のある2つ以上の構成アウトカムで構成される。複合アウトカムを構成する定義された構成アウトカムのいずれか1つを経験した参加者は、複合アウトカムを経験したとみなされる^{21,22}。複合アウトカムを用いる場合は、事前に決められ、正当化し、完全に定義しておく必要がある⁵¹。その定義には、それぞれの構成アウトカムの完全な定義と、それらをどのように組み合わせるかについての説明（例：どのような解析手順で複合アウトカムの発生を定義するか）が含まれる。

しかし、複合アウトカムは、十分に報告されている場合であっても、解釈が困難な場合がある。例えば、複合アウトカムは、構成アウトカムに対する効果が反対方向に向かう場合、または構成アウトカムの効果水準が異なる場合（例：死亡と障害を組み合わせた場合）、治療効果を隠す可能性があり、各構成要素について質の高い報告を行う必要性をさらに高めている^{22,54}。

項目 6a.7. 試験登録または試験プロトコルで事前に決められていないアウトカムを特定する。

試験プロトコルや試験登録で事前に決められていないアウトカムで、試験中に測定されたものはすべて明確に特定し、明記すべきである。事前に決められていないアウトカムは、当初予定されていなかった全く新しいアウトカムドメインの追加によって生じる可能性がある（例：病院データベースから得られた循環器病院への入院頻度の変化を予定外に

追加して解析した場合)。また、測定変量、解析指標、集計方法、評価時点が事前に決められたアウトカムと異なり、事前に決められていないアウトカムが考えられる。例えば、試験プロトコルで事前に決められた6ヵ月時点ではなく、12ヵ月時点の治療成功率が報告された場合、12ヵ月時点の治療成功率は事前に決められていないアウトカムとして特定され、具体的な変更点が説明されるべきである。アウトカムのデータの変更には多くの理由が存在する(例:スクリーニング、診断、サーベイランスの手順が変更される;コーディングシステムが変更される;新しい有害事象のデータが現れる)。主要アウトカムの基本的な変更について、試験責任医師はその詳細を報告すべきである(例:変更の性質と時期、動機、きっかけが内部データソースから生じたのか外部データソースから生じたのか、この変更を誰が提案したのか誰が承認したのか)。

試験報告書における未申告アウトカムの追加は大きな懸念事項である^{7,55)}。CONSORT承認の有力ジャーナル5誌に掲載された67試験のうち、追加されたアウトカムは365件であった(1試験あたり平均5件の未申告アウトカム)⁷⁾。追加されたアウトカムのうち、事前に決められていないとされたものは15%未満であった⁷⁾。報告されたアウトカムが試験プロトコルや試験登録の内容と一致しているかどうかの判断は、読者自身に確認させるべきではなく、負担の大きい作業であり(1試験あたり1~7時間かかる)と推定される⁷⁾、場合によっては不可能である(試験プロトコルが公開されていない場合など)。試験進行中にアウトカムを変更する正当な理由がある場合もあり、著者は報告バイアスの可能性を避けるため、このような変更を開示すべきである(CONSORT 2010 声明のチェックリスト項目6b)。

項目 6a.8. アウトカムを評価するために使用された評価尺度(例:質問票、臨床検査)を、それらの試験参加者と同様の集団における信頼性、妥当性、応答性ととも記述する。

評価尺度の十分な説明に必要とされる情報は、試験の再現性と結果の解釈可能性を確保するのに十分なものでなければならない(例:使用された評価尺

度のバージョン、実施方法、関連する検査機器のメーカーとモデルを明記する)³⁵⁾。十分な信頼性(例:再検査法、評価者間または評価者内の信頼性、内的整合性)、妥当性(例:内容妥当性、構成概念妥当性、基準妥当性、異文化間の妥当性、構造的妥当性)、評価対象である健康アウトカムの変化(すなわち、応答性)を研究手段の種類に応じて適切に検出する能力、試験参加者と類似した集団との比較が可能であることを実証できる経験的エビデンスを要約し、参考文献を示すことが不可欠である。このようなエビデンスは、測定特性に関する質の高い一次研究、評価尺度の測定特性に関するシステマティックレビュー、およびコアアウトカムセットから引き出すことができる。診断検査の精度は、定義されたアウトカムが治療前後の症状の有無に関連する際に報告に関連する⁵⁶⁾。

CONSORT-Outcomes 2022 拡張版のチェックリスト項目6a.8では、評価尺度の測定特性が異なる集団間(例:異なる年齢層間)で一般化できないと推定されているため、試験参加者と類似した集団(または少なくとも実質的に異なる集団)における関連する測定特性を記述することも推奨されている⁵⁷⁾。使われた集団に対して評価尺度の測定特性が不明な場合は、その旨をこの評価尺度を使用することが適切または必要であると考えられた理由の根拠とともに記載することができる。

試験データの質と解釈がこれらの測定特性にかかっているため、この情報を報告することは極めて重要である。例えば、内容妥当性の低い評価尺度は、測定しようとしたドメインを正確に反映しない可能性があり、評価者間の信頼性の低い評価尺度は、期待された結果が検出されなかったり、計画された検出力計算で説明されなかったりした場合、試験の統計的検出力³⁵⁾を損なう可能性もある³⁰⁻³²⁾。

項目 6a.9. アウトカムを評価した人(例:看護師、保護者)と、アウトカムを評価するために評価尺度の実施に必要な資格または試験に特化した訓練を記述する。

多くの種類のアウトカム(例:行動的アウトカム、心理的アウトカム)において、関心のあるアウトカ

ムを誰が評価するかによって、かなり異なった回答、ひいては異なった試験結果が得られる。このようなばらつきは、評価者の訓練や経験の違い、視点の違い、患者の想起 (recall) から生じる可能性がある^{58,59}。臨床医、患者、非臨床オブザーバーから報告された臨床アウトカムの評価、またはパフォーマンスに基づく評価は、米国食品医薬品局 (US Food and Drug Administration) によって、それぞれに臨床医報告アウトカム、患者報告アウトカム、オブザーバー報告アウトカム、パフォーマンスアウトカムに分類されている⁶⁰。

様々な人によって評価されうるアウトカムについては、試験の観点から行われたアウトカム評価者の選択に関する説明を行うべきである。アウトカム評価者の影響を受けないアウトカム (例: 血漿コレステロールの水準) については、この情報はあまり関連がない。試験担当者がアウトカム評価者として機能するために必要な専門資格または試験に特化した訓練は、多くの場合、説明に関連している³⁵ (例: ウェクスラー式知能検査 (Wechsler Abbreviated Scale of Intelligence) 第2版を使用する場合、博士号または心理学博士号を有し、関連する患者集団について5年以上の経験を持ち、本尺度または類似のIQ評価法を過去15回以上実施したことのある評価者が必要な場合がある)。患者の治療割付けや新たな試験結果に対する評価者の盲検化に関する詳細は、CONSORT 2010 声明のチェックリスト項目 11a に記載されている。

項目 6a.10. データ取集中 (例: 重複測定) およびデータ収集後 (例: アウトカムデータの値の範囲検査) にアウトカムデータの質を高めるために使用されたプロセスを記述する、またはそれらの詳細を確認できる場所を記載する。

臨床試験のデータ取集中および収集後に、アウトカムデータの質を高めるために使用されたプロセスの説明を示すことは、透明性を担保し、試験データの質の評価を容易にする。例えば、主観的アウトカム評価が重複して実施された場合 (例: 病理性的評価) や、独立かつ正確なアウトカム評価を保証するために中央判定委員会 (central adjudication committee) が使用された場合が挙げられる。その

他の一般的な例としては、データが適切な形式 (例: 整数) であること、データが予想される値の範囲内であること、独立した直接閲覧 (source document verification) (例: 外部の臨床開発モニターによる) が行われていることなどがある³⁵。試験報告書には、完全な説明、または完全な情報を入手できる場所 (例: オープンアクセスの試験プロトコル) に関する簡単な要約を含めるべきである。

サンプルサイズの記述に関する CONSORT-Outcomes 2022 拡張版のチェックリスト項目

項目 7a.1. 治療群間の目標差 (例: 最小重要差) を定義し、正当化する。

この項目は、CONSORT 2010 声明のチェックリスト項目 7a である「どのように目標症例数が決められたか」を報告することを拡張したもので、試験のサンプルサイズ算出に使用した目標差を報告するよう著者に促すものである。目標差は、1つ以上の主要な利害関係者グループによって現実的または重要であるとみなされるべき解析において、特定の時点に介入群間の主要アウトカムで検出しようとする差として、サンプルサイズの計算に使用される値である³⁷。The Difference Elicitation in Trials project は、試験の目標差の選択、サンプルサイズの計算、報告に関する汎用的なエビデンスに基づくガイダンスを公表している^{37,38}。目標差は、最小重要差 (MID; 重要であると捉えられる患者間の最小の差)^{24,26,27} または最小有益効果 (患者によって決められた介入のコスト、潜在的な害、不便さを正当化できる、介入による最小の有益な効果) とする可能性がある³³。選択された目標差に情報を示すには、実地的 (pragmatic) または臨床的な要因が様々であり (例: 主要アウトカムを評価するために使用された評価尺度における信頼できる MID の利用可能性)⁴⁷、また、多数の異なる選択肢が利用可能であるため (例: 複数の MID のうちの1つ、またはパイロット研究に基づく値)⁴⁷、選択された目標差がなぜ選ばれたのかを説明することが重要である^{23,48,49}。

統計的手法の記述に関する CONSORT-Outcomes 2022 拡張版のチェックリスト項目

項目 12a.1. 主要および副次的アウトカムの解析または解釈における多重性を説明するために使用した

手法を記述する（例：共主要アウトカム（coprimary outcomes）、複数の時点で評価された同一アウトカム、アウトカムのサブグループ解析）。

この項目は、CONSORT 2010 声明のチェックリスト項目 12a を拡張したもので、結果の解析や解釈に関連する多重性を説明するために使用した統計的手法を記述するよう著者に促すものである。アウトカムの多重性問題は、臨床試験において一般的であり、特に注意すべきなのは、共主要アウトカムがある場合、一つのアウトカムを繰り返し評価することによって複数の評価時点が存在する場合、一つのアウトカムについて複数の解析が計画されている場合（例：中間解析やサブグループ解析、多群間比較試験）、解析対象となる副次的アウトカムが多数ある場合である⁶¹⁾。

このような多重性を説明するために使用される方法には、統計的手法（例えば、ファミリーワイズエラー率（family-wise error rate））や記述的アプローチ（例：探索的解析であることを指摘し、予想された偽陽性のアウトカムの数を考慮して結果を述べる）が考えられる^{61,62)}。このような情報は、報告書の本文に簡潔に記述することもできるし、統計解析計画書に詳しく記述することもできる⁶³⁾。多重性を説明する方法が使用されなかった場合、著者はその旨を報告することができる（例：該当しない場合、必要ないと判断された場合）。

項目 12a.2. 解析および報告からアウトカムデータを除外する基準を記載し、正当化する、またはアウトカムデータを除外しなかったことを報告する。

この項目は、CONSORT 2010 声明のチェックリスト項目 12a を拡張したもので、著者に（1）特定のアウトカムデータを解析から除外するために適用した基準を記載し、正当化すること、または（2）アウトカムデータを除外しなかったことを報告することを推奨する。これは、明示的かつ意図的に除外されたアウトカムデータに関するものであり、例えば、参加者が記入済みの質問票に欠測項目が多すぎる場合や、特定のアウトカムの外れ値を正当に除外した場合がこれに該当する。これは、読者が報告された結果を解釈するのに役立つ。この情報を、治療群ごとにそれぞれのアウトカムデータの除外理由が

記載されている CONSORT フローチャートに示すことができる。

項目 12a.3. 欠測のパターン（例：ランダムでない欠測（missing not at random））を評価するために使用された方法、およびアウトカム項目または評価全体の欠測を対処するために使用された方法を記述する。

試験において参加者のアウトカムデータの欠測はよく見られることであり、多くの場合、この欠測はランダムではなく、治療群への割付け、患者特異的（予後）因子、または特定の健康アウトカムの発生のいずれかに関連していることを意味する^{64,65)}。欠測データがある場合、CONSORT-Outcomes 2022 拡張版のチェックリスト項目 12a.3 では、（1）欠測のパターンを評価または特定するために用いた方法、および（2）統計解析においてアウトカムまたは評価全体の欠測を対処するために用いた方法（その方法の選択は、特定された欠測パターンによって決定される）を記述することを推奨している（例：多重代入法（multiple imputation）、完全ケース分析（complete case）、尤度（likelihood）に基づく方法、逆確率重み付け（inverse-probability weighting））。

欠測データは、そのパターンによって、試験の検出力低下につながり、結論に影響を与え、試験のバイアスリスクを左右するため、欠測データのパターンと対処法について情報を提供することは非常に重要である^{66,67)}。欠測の程度や解析における欠測データの扱い方が明確でないと、メタ解析者が統合推定値（pooled estimates）で重み付けをするために必要なサンプルサイズを正確に抽出することができなくなり、報告された結果における欠測データから生じるバイアスのリスクを正確に評価することもできなくなる^{67,68)}。このチェックリスト項目は、完全なデータセットがある場合には適用されず、欠測データの量が無視できると考えられる場合には重要でないかもしれない。

欠測のパターン（欠測データのメカニズムとも呼ばれる）には、完全にランダムな欠測（missing completely at random）、ランダムな欠測（missing at random）、ランダムでない欠測があり、読者やメタ解析者がデータセットにどのようなパターンが存在

するかを判断するのに役立つよう、試験での記述が必要である⁶⁹⁾。欠測データの一部は、ランダムな欠測かどうかを見分けるために、まだ測定できるかもしれない(例:欠測データのある試験参加者のサブセットに対する協調的な追跡調査によって測定する)⁷⁰⁾。欠測のパターンは、欠測したアウトカムや評価全体の対処法の選択に関連しており(例:多重代入法や最尤法(maximum likelihood)解析は、データが少なくともランダムな欠測であると仮定している)、報告することが不可欠である。試験結果の頑健性を評価するために実施された感度分析(例:欠損データを処理するために様々な方法を用いる)はすべて報告されるべきである^{35,64)}。

項目 12a.4. 試験プロトコルの不遵守に関連するアウトカム解析集団の定義を示す(例:ランダム化解析として)。

臨床試験のアウトカムデータは、様々な方法で解析され、異なる結果を導き出す可能性がある。CONSORT 2010 声明に記載された一般的な報告原則は、「有識者である読者が元のデータにアクセスでき、報告された結果を検証できるように、統計的手法を十分に詳細に記述する」ことであった²⁾。項目 12a.4 は、CONSORT 2010 声明のチェックリスト項目 12a を拡張したもので、著者が試験プロトコルの不遵守に関連する各解析で使用したアウトカム解析集団の定義を報告することを推奨している。各解析において、試験責任医師が最初に割付けられた群にランダムに振り分けられた全ての参加者を含めたかどうか(すなわち、intention-to-treat 解析)の情報は、臨床試験結果の批判的評価と解釈にとって非常に重要であることが広く認識されている^{2,35)}。

欠測データの量はアウトカムによって異なり、欠測の理由も様々であるため、CONSORT-Outcomes 2022 拡張版のチェックリスト項目 12a.4 では、統計解析に使用したアウトカム解析集団の定義を報告することを明記している。例えば、ある完全なデータセットが死亡率のアウトカム解析に利用できるが、同じ試験内で患者報告アウトカムの解析に利用できない場合がある。別の例では、害の解析は、害の発生の有無が一度も受けていない治療に起因することがないように、試験の介入を受けた参加者に限定さ

れる可能性がある³⁵⁾。

アウトカムと推定の記述に関する CONSORT-Outcomes 2022 拡張版のチェックリスト項目

項目 17a.1. 事前に決められたすべてのアウトカム解析の結果を含める、または本報告書に記載がない場合はどこに結果の記載があるかを示す。

この項目は、アウトカムと推定に関する CONSORT 2010 声明のチェックリスト項目 17a を拡張したもので、著者に試験プロトコルや統計解析計画書で事前に決められたすべてのアウトカム解析の結果を報告していることを確認するよう注意を促すものである⁶⁸⁾。これは標準的な実施方法であると期待されるが²⁾、読者が試験結果の選択的不報告があったかどうかを判断するにあたって、試験報告書に記載されている事前に決められた解析に関する情報が不十分であることはしばしばある⁷¹⁾。1回の試験報告書で事前に決められたすべての解析を報告することが不可能な場合(例:事前に決められた副次的アウトカムが多数ある試験)、著者は、事前に決められたその他のアウトカム解析の結果がどこにあるかを報告するべきであり(例:リンクされた出版物やオンラインリポジトリ)、または長期追跡の場合、後で報告する意向を示すべきである。

補助的分析の記述に関する CONSORT-outcomes 2022 拡張版のチェックリスト項目

項目 18.1. 事前に決められていない解析があった場合は、その理由を説明する。

アドヒアランス(adherence)に関する最近の研究では、出版された臨床試験において、事前に統計解析が決められることが少ないままであり、事前に決められた解析と報告された解析の間に説明できない不一致があることが示された⁷¹⁾。この項目は、CONSORT 2010 声明の補助的分析に関するチェックリスト項目 18 を拡張したもので、事前に決められていないが(例:試験プロトコルや統計解析計画書で)、試験報告書で報告されたすべての解析について説明を行う必要があることを推奨している。このような解析は、探索的解析と、または事前に決められていない解析とも呼ばれる。事前に決められていない解析が実施・報告された根拠を示すことは、試験の透明性を高め、試験の信頼性を正しく評価す

るために重要である。そのような追加解析がいつ行われたかを示すことも重要である（例：他のアウトカムの比較解析結果を確認する前、またはその後）。同じデータを複数回解析すると偽陽性の結果が出るリスクがあり、事前に決められていない解析を選択的に報告することによってバイアスが生じる可能性もある。

考察

CONSORT-Outcomes 2022 拡張版は、エビデンスおよびコンセンサスに基づき、出版された臨床試験報告書のアウトカム報告に関するガイダンスを示すものであり、CONSORT 2010 声明のチェックリストに17の報告項目を追加し、報告に関する推奨事項をSPIRIT-Outcomes 2022 拡張版のガイダンスと調和させたものである¹⁷⁾。これら2つの拡張版ガイドラインを整合させることで、試験プロトコルから試験完了までの一貫した連続的な報告が行われ、研究者による試験プロトコルと試験報告書の作成が円滑になり、さらに重要なこととして、最終報告書における試験プロトコル遵守の評価も容易になる²⁰⁾。CONSORT 2010 声明と同様に⁴¹⁾、CONSORT-Outcomes 2022 拡張版は、表や図、オンラインのみの補足資料を含む試験報告書の内容に適用される⁷²⁾。現行の推奨事項も同様に、これらの情報を報告する構成や場所については規定していない。著者は、論文中のどこかで、チェックリストの項目を、十分に詳細かつ明瞭に記載すべきである⁴¹⁾。

CONSORT-Outcomes 2022 拡張版のチェックリストのユーザーは、これらの追加チェックリスト項目がアウトカム報告に最低限必要な項目であり、試験の有用性、透明性、再現性を最大化し、結果の選択的不報告を制限するために、CONSORT 2010 声明^{2,3)}のガイドラインに追加されたものであることに留意すべきである（付録のeTable 5）。場合によっては、アウトカム特有の追加情報を試験報告書で報告することが重要である⁴⁾。その場合、付録のeTable 6にあるようなもの、PROに特化した報告ガイダンスを示すCONSORT-PRO¹⁴⁾や害の報告に関するCONSORTの拡張版¹³⁾を参照するのが考えられる。CONSORT-Outcomes 2022 拡張版

に準拠する著者は、ある項目が自分の試験に関連しない理由を説明する必要がある。例えば、この拡張版チェックリストは、系統的に評価されたアウトカムを報告するためのものであるが、自発的に報告された有害事象のような系統的に収集されていない、または事前に決められていないアウトカムには適用できない可能性がある。ジャーナルの文字数に制約がある場合、著者は、オープンアクセスの試験プロトコル、統計解析計画書、試験登録情報を参照し、オンラインのみの補足資料を示すことができる。

CONSORT-Outcomes 2022 拡張版の主なユーザーは、試験論文の著者、ジャーナル編集者、査読者、システマティックレビュー作成者、メタアナリシス研究者、学術機関、患者（試験参加者を含む）、一般市民であると予想される（付録のeTable 5）。これらのグループがこの拡張版を使用することで、試験の有用性、透明性、再現性の向上に役立つ可能性がある。患者および一般市民の参加は、方法論的に複雑なトピックが議論されたコンセンサス会議にうまく組み込まれており、これは、これまでの報告ガイドライン作成において稀なことであった。今後の報告ガイドラインの作成には、全工程を通して患者や一般市民を参加させるべきである。CONSORT-Outcomes 2022 拡張版を、EQUATOR ネットワークやCONSORTのウェブサイトを含むところで、先述したように普及させる予定である¹⁵⁾。エンドユーザーは、内容、わかりやすさ、使いやすさについてオンラインで意見を提供することができ⁷³⁾、今後のアップデートにつながられる。

限界

本研究にはいくつかの限界がある。第1に、本チェックリストに含まれている項目は、ほとんどの潜在的な利点と一部の害を含む系統的に収集されたアウトカムに適している。しかし、系統的に評価されていない害の報告には、他の項目が適用できる可能性がある⁷⁴⁾。

第2に、これらのチェックリスト項目は、まだCONSORTの主要チェックリストに統合されていない

いため、著者や編集者によっては、複数のチェックリストを確認して使用することが負担になると考えられ、それが利用率に影響する可能性がある⁷⁵⁾。今後、これらの追加チェックリスト項目をCONSORTの主要チェックリストに統合する取り組みにより、実際の運用が促される可能性がある。

第3に、より広範な研究コミュニティにおける使いやすさを向上させる目的で、専門家およびエンドユーザーによる大規模な多国籍研究グループがこれらの推奨事項の作成に関与したにもかかわらず、パネリストは自己選択的であったため（すなわち、関心のある個人がデルファイ調査に参加するために署名を行った）、デルファイ調査の投票結果は無回答バイアスの影響を受けた可能性がある。

第4に、コンセンサス会議のパネリストは、彼らの臨床試験の実施、監督、報告に関連する専門知識

と役割に基づき、意図に抽出された¹⁵⁾。コンセンサス会議のパネリストに十分に代表されていない人々（例：北米や欧州以外の試験担当者）の見解は異なる可能性がある。アウトカム報告のガイダンスに関する厳密なスコーピングレビュー^{4,18)}を含め、本ガイドラインの作成に用いられた系統的でエビデンスに基づいたアプローチは^{15,16)}、これらの限界の潜在的な影響を和らげるのに役立つかもしれない。

結論

本 CONSORT 2010 声明の CONSORT-Outcomes 2022 拡張版は、すべての出版された臨床試験報告書で検討すべき17のアウトカム特有の報告項目を提供し、試験の有用性、再現性、透明性の向上に役立ち、試験結果の選択的不報告のリスクを最小限に抑える可能性がある。

参考文献

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, SpernaWeiland 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://faculty.ksu.edu.sa/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/CBO9780511996214
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): 200207. doi:10.7326/0003-4819-158-3-20130205000583
37. Cook JA, Julious SA, Sones W, et al. Practical help for specifying the target difference in sample size calculations for RCTs: the DELTA2 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. DELTA2 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/NEJMs1012065
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; DeCIde 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. Mouelhi 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
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**: 282282. 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. *JBIM Manual for Evidence Synthesis*. JBI; 2020. doi:10.46658/JBIMES-20-10
57. Mookink 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
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
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**: 186196. 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