

Guidelines for Reporting Outcomes in Trial Protocols The SPIRIT-Outcomes 2022 Extension の和訳

著者 Nancy J. Butcher^{1,2}, Andrea Monsour¹, Emma J. Mew^{1,3}, An-Wen Chan⁴, David Moher^{5,6}, Evan Mayo-Wilson⁷, Caroline B. Terwee^{8,9}, Alyssandra Chee-A-Tow¹, Ami Baba¹, Frank Gavin¹⁰, Jeremy M. Grimshaw^{11,12}, Lauren E. Kelly^{13,14}, Leena Saeed¹, Lehana Thabane¹⁵, Lisa Askie¹⁶, Maureen Smith¹⁷, Mufiza Farid-Kapadia¹, Paula R. Williamson¹⁸, Peter Szatmari^{19,20}, Peter Tugwell^{6,12,21,22}, Robert M Golub²³, Suneeta Monga^{2,20}, Sunita Vohra²⁴, Susan Marlin^{25,26}, Wendy J Ungar^{1,27}, Martin Offringa^{1,27,28}

翻訳 馬場 亜沙美 (BABA Asami)^{1*}, 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, 新林 史悠 (SHINBAYASHI Fumiharu)¹, 板橋 怜央 (ITABASHI Reo)¹, 柿沼 俊光 (KAKINUMA Toshihiro)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, アウトカム, SPIRIT

Guidelines for Reporting Outcomes in Trial Protocols The SPIRIT-Outcomes 2022 Extension

Authors:

Nancy J. Butcher^{1,2}, Andrea Monsour¹, Emma J. Mew^{1,3}, An-Wen Chan⁴, David Moher^{5,6}, Evan Mayo-Wilson⁷, Caroline B. Terwee^{8,9}, Alyssandra Chee-A-Tow¹, Ami Baba¹, Frank Gavin¹⁰, Jeremy M. Grimshaw^{11,12}, Lauren E. Kelly^{13,14}, Leena Saeed¹, Lehana Thabane¹⁵, Lisa Askie¹⁶, Maureen Smith¹⁷, Mufiza Farid-Kapadia¹, Paula R. Williamson¹⁸, Peter Szatmari^{19,20}, Peter Tugwell^{6,12,21,22}, Robert M Golub²³, Suneeta Monga^{2,20}, Sunita Vohra²⁴, Susan Marlin^{25,26}, Wendy J Ungar^{1,27}, Martin Offringa^{1,27,28}

Translators:

Asami Baba^{1*}, Naoko Suzuki¹, Kazuhiko Noda¹, Eri Hatano¹, Noriyuki Takahashi¹, Fumiharu Shinbayashi¹, Reo Itabashi¹, Toshihiro Kakinuma¹, Kazuo Yamamoto¹

*Correspondence author: Asami Baba

Keywords: Human clinical trial, Human trials, Food for specified health uses, Food with function claims, Research design, Guideline, Outcome, SPIRIT

Affiliations (Authors):

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

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

* 責任著者: 馬場 亜沙美

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

- 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 Protocols The SPIRIT-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.2022.21243) にアクセスして参照する。

要約 [Abstract]

重要性 [IMPORTANCE]

試験プロトコルに試験結果に関する完全な情報を記載することは、規制当局の承認を得るため、標準化された試験実施を確保するため、研究の無駄を省くため、そして試験方法の透明性を提供し、試験の再現、批判的評価、試験結果の正確な報告と解釈、知識の統合を促進するために極めて重要である。しかし、どのようなアウトカム特異的情報を含めるべきかについての推奨は多様で一貫性がない。透明で再現性のあるアウトカムの選択、評価、解析を促進する報告方法を改善するためには、臨床試験プロトコルにどのようなアウトカム特異的情報を記載すべきかについて、具体的で統一されたガイダンスが必要である。

目的 [OBJECTIVE]

標準プロトコル項目と統合することにより、臨床試験プロトコルにおけるアウトカムを記述するための、エビデンスとコンセンサスに基づく統一された基準を開発する：Recommendations for Interventional Trials (SPIRIT) 2013 声明と統合する。

エビデンスレビュー【EVIDENCE REVIEW】

Enhancing the Quality and Transparency of Health Research (EQUATOR) の方法論的枠組みを用いて、SPIRIT 2013 声明の拡張版である SPIRIT-Outcomes 2022 は、下記の通り作成された。

- (1) 専門家との協議によるアウトカム報告項目の候補の作成と評価、専門家の募集、MEDLINE および Cochrane Methodology Register 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وپングレビューにより作成された。
- (2) 22カ国124人のパネリストによる3回の国際デルファイ投票プロセス（2018年11月～2019年2月）を実施し、追加項目の評価と特定を行った。
- (3) 25人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）を実施し、臨床試験プロトコルで扱うべきアウトカム特異的報告の必須項目を特定した。

調査結果【FINDINGS】

スコوپングレビューおよび専門家との協議により、臨床試験プロトコルで扱うべきアウトカム特異的報告に関連する108の推奨事項が特定され、その大部分（72%）はSPIRIT 2013 声明には含まれていなかった。デルファイ調査の結果、19の項目がコンセンサス会議での更なる評価基準を満たし、SPIRIT-Outcomes 2022 拡張版に含めることが可能となった。コンセンサス会議での議論の結果、SPIRIT 2013 声明チェックリスト項目をさらに詳細にした9項目が得られた。この9項目は、臨床試験プロトコルにおいて、主要アウトカム、副次的アウトカム、その他のアウトカムの選択を完全に定義し、正当化すること（SPIRIT 2013 声明チェックリスト項目¹²⁾、サンプルサイズの計算に使用する主要アウトカムの治療群間目標差を定義し、正当化すること（SPIRIT 2013 声明チェックリスト項目¹⁴⁾、アウトカムを評価するために使用する評価尺度の応答性を記述し、アウトカム評価者についての詳細を提供すること（SPIRIT 2013 声明チェックリスト項目18a）、解析または結果の解釈に関連する多重性を考慮するための計画された方法を記述すること（SPIRIT 2013 声明チェックリスト項目20a）に関するものである。

結論と関連性【CONCLUSIONS AND RELEVANCE】

SPIRIT-Outcomes 2022 は、SPIRIT 2013 声明を拡張したもので、すべての臨床試験プロトコルで扱うべき9つのアウトカム特異的報告項目を示しており、臨床試験の有用性、再現性、透明性の向上に役立ち、試験結果の選択的非報告のリスクを最小限に抑える可能性がある。

要点【Key Points】

問題【Question】

臨床試験プロトコルにはどのようなアウトカム情報を含めるべきか？

結果【Findings】

Enhancing the Quality and Transparency of Health Research (EQUATOR) の方法論的枠組みの手法を用いた、エビデンスおよび国際的なコンセンサスに基づくアプローチにより、臨床試験プロトコルに記載すべき9つのアウトカム特異的報告項目が特定された。

意義【Meaning】

臨床試験プロトコルにこれらの項目を含めることで、試験の有用性、再現性、透明性が向上し、試験結果の選択的非報告を抑制することに役立つであろう。

試験プロトコルには、ランダム化臨床試験の目的、デザイン、方法、計画された解析、組織、修正事項が記載されている。試験プロトコルは、試験責任医師、試験スタッフ、資金提供機関、規制機関、医療技術評価機関、倫理審査委員会、システマティックレビュー、学術雑誌などで使用される。標準プロトコル項目 Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2013 声明¹⁻³⁾ は、臨床試験プロトコルに含めるべき 33 の報告項目のチェックリストを提供した⁴⁾。患者報告アウトカムを報告するための SPIRIT 拡張版が利用可能であるにもかかわらず⁵⁾、系統的に収集されたすべてのアウトカムの種類、集団、および試験デザインに適用可能なアウトカム特異的情報を報告するためのエビデンスに基づくガイドラインは存在しない。

試験プロトコルでアウトカム特異的情報を完全に報告することは、倫理的および規制当局の承認を得るために重要であり、また、試験が事前に設定された目的と方法に従って実施されていることを保証するものである。試験プロトコルが明確に記述されていれば、個々の試験の方法の透明性が促進され、試験結果の報告や解釈が容易になるが、最近のスコーピングレビュー⁶⁾では、学術的、規制的、公的な情報源による試験プロトコルにおけるアウトカムの選択、評価、分析を前向きに報告する方法に関する勧告は多様で、一貫性がなく、多くの文書に分散していることが明らかになった。SPIRIT 2013 声明^{1,2)}がアウトカムの報告方法に関する一般的なガイダンスを提供しているにもかかわらず、アウトカムの報告に関する問題は依然として残っている⁷⁻¹¹⁾。このような報告の問題は、システマティックレビューやメタアナリシスの結論に影響を及ぼし、現在進行中の研究の無駄の一因となっている^{12,13)}。

SPIRIT-Outcomes 2022 拡張版の目的は、試験プロトコルのための、調和された、エビデンスとコンセンサスに基づくアウトカム報告基準を開発することであった。

方法【Methods】

SPIRIT-Outcomes 2022 拡張版は、Instrument for Reporting Planned Endpoints in Clinical Trials (InsPECT) プロジェクト¹⁴⁾の一環として、ガイドライン開発報告のための Enhancing Quality and Transparency of

Health Research (EQUATOR) の方法論的枠組みに従って開発された¹⁵⁾。The Hospital for Sick Children の研究倫理委員会の決定により、倫理承認は不要であった。統計解析計画の内容に関する具体的なガイダンスが公表されている¹⁶⁾。SPIRIT-Outcomes 2022 拡張版の開発は、臨床試験報告のための Consolidated Standards of Reporting Trials (CONSORT)-Outcomes 2022 拡張版と並行して行われた¹⁷⁾。

主な開発段階【Key Development Phases】

第 1 に、専門家との協議や、専門家の提案、MEDLINE と Cochrane Methodology Register の電子データベース検索、灰色文献検索、および記載された参考文献リスト検索によって特定されたアウトカム報告に関する既存のガイダンス (2018 年 3 月 19 日以前の 10 年間に発表されたもの) のスコーピングレビューから統合された、ランダム化臨床試験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^{6,18)}。第 2 に、2018 年 11 月~2019 年 2 月にかけて 3 回にわたる国際的なデルファイ投票プロセスを実施し、追加項目を特定し、9 段階のリッカー尺度を用いて各項目の重要性を評価し、22 カ国 124 人のパネリストが記入した (付録の eTable 1 ※)。第 3 に、患者パートナーおよび公的パートナーを含む 4 カ国 25 人のパネリストが参加した対面専門家コンセンサス会議 (2019 年 4 月 9 日~10 日) を開催し、試験プロトコルにおけるアウトカム特異的情報の報告に関連する必須項目リストを特定し、普及活動を確立した。項目の選択と文言は、コンセンサス会議後の幹部パネリストによる会議と、コンセンサス会議のパネリストとの電子メールを通じて最終決定された。

その他の情報【Other Information】

SPIRIT Outcomes 2022 拡張版の開発に関する詳細な方法は、各段階での評価項目数、最終的な対象項目セットへのプロセスを含め、付録の eAppendix 1 ※に掲載されている (付録の eFigure ※)。スコーピングレビュー試験プロトコルと結果は公表されており^{6,18)}、付録の eAppendix 1 ※に、検索戦略は付録の eAppendix 2 ※に掲載されている。デルファイ投票のパネリストと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 ※に掲

載されている。患者パートナーと公的パートナーの関与に関する詳細は、付録の eAppendix 1 ※に掲載されている。

結果【Results】

SPIRIT 2013 声明のチェックリスト項目を含めることに加え、SPIRIT-Outcomes 2022 拡張版では、試験デザインや集団にかかわらず、最低9つのアウトカム特異的報告項目の記述を試験プロトコルに前向きに含めることを推奨している。スコーピングレビューと専門家との協議により、試験プロトコルに記載すべきアウトカム特異的報告に関連する108の推奨事項が特定されたが、その大部分(72%)はSPIRIT 2013 声明には含まれていなかった。すべての推奨事項は、デルファイ投票のために56項目に統合され、デルファイ調査の結果、19項目がコンセンサス会議での更なる評価基準を満たし、SPIRIT-Outcomes 2022 拡張版に含めることが可能となった。SPIRIT 2013 声明チェックリスト項目と、SPIRIT-Outcomes 2022 拡張版で追加された9つのアウトカム特異的報告項目を表1に示す¹⁹⁾。

チェックリストの記入可能なバージョンは、付録の表3~4 ※およびSPIRIT Web サイトに掲載されている³⁾。更新されたチェックリストを使用する場合、ユーザーは用語集²⁰⁻³⁷⁾ (Box) で重要な用語の定義を参照する必要がある。定義は分野や地理的領域を超えて存在する。定義された結果の5つの中心要素(例付き)を表2に示す^{38,39)}。

SPIRIT-Outcomes 2022 拡張版の新しいチェックリスト項目をSPIRIT 2013 声明と合わせて適用することにより、アウトカムが試験プロトコルで包括的に前向きに定義され、試験報告書で報告されることが保証される。9つのSPIRIT-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、彼らの提案する行動およびその結果もたらされる潜在的な利益の推定リストは、付録の eTable 5 ※に掲載されており、コンセンサス会議の学術的解釈会合により作成された。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、試験の著者、雑誌編集者、査読者、システムティックレビュー担当者、患者、一般市民、試験参加者にとって有用であろう(付録の eTable 5 ※)。

本報告書には、SPIRIT-Outcomes 2022 拡張版か

ら作られた9つ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のSPIRIT 2013 声明¹⁾、および解説ガイドライン報告書に記載されている²⁾。一部の試験プロトコルまたは関連する試験文書(統計解析計画書など)に含めると有用な追加項目は、付録の eTable 6 ※に記載されているが、すべての臨床試験プロトコルで必須の報告項目とは考えられていない。

Box. SPIRIT-Outcomes 2022 拡張版で使用される用語集【Glossary of Terms Used in the SPIRIT-Outcomes 2022 Extension】

複合アウトカム【Composite outcome】：複合アウトカムは、2つの要素アウトカム(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。構成要素で指定された事象のいずれか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^{20,21)}。

CONSORT 2010：2010年に発表された Consolidated Standards of Reporting Trials (CONSORT) 声明^{22,23)}。

CONSORT-Outcomes 2022 拡張版【CONSORT-Outcomes 2022 extension】：CONSORT2010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必須チェックリストの追加項目。

構成概念妥当性【Construct validity】：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³³⁾。

基準妥当性【Criterion validity】：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³³⁾。

最小重要変化量【Minimal important change】：患者、臨床医、その他関係者が重要と考える、患者内での最小の変化量^{24,25)}。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

(二分法アウトカム) の変化量である。この用語は、医療に関する文献では、しばしば minimal important difference (最小重要差) という用語と互換的に使用される。SPIRIT-Outcomes 2022 拡張版では、最小重要変化量は概念的に重要な患者内変化量 (項目 12.2) を指し、最小重要差は重要な群間差を指す。minimal の代わりに minimum を用いたり、clinically や clinical という形容詞を付けたりするなどの細かい用語の変形がよく見られる (例: the minimum clinically important change) ²⁶⁾。

最小重要差【Minimal important difference】 : 患者、臨床医、その他関係者が重要であると考えた最小の群間差 ^{24, 27-29)}。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 (二分法アウトカム) の場合がある。minimal の代わりに minimum を用いたり、clinically または clinical という形容詞を加えたりするなどの、細かい用語の変形はよく見られる (例: the minimum clinically important difference) ²⁶⁾。

アウトカム【Outcome】 : 健康介入への曝露の効果を調べるために評価されるものを指す ³⁰⁾。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【Primary outcome】 : 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 ⁵⁾。これは通常、効果 (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アウトカムである ³¹⁾。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 ⁶⁾。

応答性【Responsiveness】 : アウトカムドメインにおける経時的変化を正確に検出し測定するための評価尺度の能力 ^{25, 32)}。単一のスコアの妥当性に言及する調査票の構成概念妥当性や基準妥当性とは異なり、反応性は変化スコアの妥当性 (すなわち縦断的妥当性) を指す ³³⁾。

副次的アウトカム【Secondary outcomes】 : 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 ⁵⁾。

最小有益効果【Smallest worthwhile effect】 : 患者によって決められた介入のコスト、潜在的な有害性、不便さを正当化できる、介入による最小の有益な効果 ³⁴⁾。

SPIRIT 2013: 2013 年に発表された Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 声明 ^{1, 2)}。

SPIRIT-Outcomes 2022 拡張版【SPIRIT-Outcomes 2022 extension】 : 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。

評価尺度【Study instrument】 : 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 ³⁵⁾。

目標差【Target difference】 : 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 ^{36, 37)}。

妥当性【Validity】 : 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 ³³⁾。

アウトカム記述のための SPIRIT-Outcomes 2022 拡張版チェックリスト項目【SPIRIT-Outcomes 2022 Extension Checklist Items for the Descriptions of the Outcomes】

項目 12.1. 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 SPIRIT 2013 声明チェックリスト項目 12 を拡張したもので、著者は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 (臨床的

表 1 SPIRIT 2013 声明および SPIRIT-Outcomes 2022 拡張版より、試験プロトコルで扱うべき推奨チェックリスト項目^a

章 (Section)	項目番号 (Item No.)	SPIRIT 2013 声明 (SPIRIT 2013 statement)	項目番号 (Item No.)	SPIRIT-Outcomes 2022 拡張版 (SPIRIT-Outcomes 2022 extension)
行政情報 [Administrative information]				
タイトル	1	試験デザイン, 集団, 介入, および該当する場合は試験の略称を示す記述的タイトル		
試験登録	2a	臨床試験の識別子と登録名。未登録の場合は, 登録予定の登録機関名		
	2b	世界保健機関 (WHO) 試験登録データセットからの全項目		
プロトコルバージョン	3	日付とバージョン識別子		
資金	4	財政的, 物質的, その他の支援源とその種類		
役割と責任	5a	プロトコル作成者の氏名, 所属, 役割		
	5b	試験依頼者の氏名と連絡先		
	5c	研究計画, データの収集・管理・分析・解釈, 報告書の作成, および報告書の出版に向けた提出の決定における, 研究スポンサーおよび資金提供者がいる場合はその役割 (これらの活動に対する最終的な権限を有するかどうかも含む)		
	5d	調整センター, 運営委員会, エンドポイント判定委員会, データ管理チーム, その他試験を監督する個人またはグループ (該当する場合) の構成, 役割, 責任 (データモニタリング委員会については項目 21a を参照)		
はじめに [Introduction]				
背景と目的 [Background and objectives]	6a	各介入の有益性と有害性を検討した関連研究 (公表および未公表) の要約を含む, 研究課題と試験を実施する正当性の記述		
	6b	比較対象の選択に関する説明		
目的	7	具体的な目的または仮説		
試験デザイン	8	試験の種類 (例: 並行群, クロスオーバー, 要因, 単一群), 割付比率, 枠組み (例: 優越性, 同等性, 非劣性, 探索的) を含む試験デザインの記述		
方法: 参加者, 介入, アウトカム [Methods: participants, interventions, and outcomes]				
研究設定	9	研究設定 (地域診療所, 大学病院など) およびデータ収集国のリストの記述。試験実施施設のリストが入手できる場所への言及		
適格基準	10	参加者の選択基準および除外基準。該当する場合は, 研究施設および介入を実施する個人の適格基準 (例: 外科医, 心理療法士, 臨床心理士など)		
介入	11a	いつ, どのように実施するかを含め, 再現可能な各グループの介入の十分な詳細 (具体的なガイダンスについては, TIDieR チェックリストおよびガイドを参照) ¹⁹⁾		
	11b	特定の試験参加者に対して割り付けられた介入を中止または変更する基準 (例: 有害性, 参加者の要求, または疾患の改善/悪化に応じた薬物投与量の変更)		
	11c	介入プロトコルの遵守を向上させるための戦略, および遵守をモニタリングするための手順 (例: 錠剤の返却, 臨床検査)		
	11d	関連する併用治療および試験中に許可または禁止される介入		

表 1-2

アウトカム	12	特定の測定変数 (例: 収縮期血圧), 解析指標 (例: ベースラインからの変化量, 最終値, イベント発生までの時間), 集計方法 (例: 中央値, 割合), 各アウトカムの時点を含む主要アウトカム, 副次的アウトカム, その他のアウトカム。選択した有効性と有害性のアウトカムの臨床的関連性の説明を強く推奨	12.1	試験の主要アウトカムにドメインを選択した根拠を提示
			12.2	主要アウトカムの分析指標が参加者内の変化量を表す場合, 個人における最小重要変化量の定義と正当化
			12.3	収集した結果データが連続的であるが, カテゴリー分析 (集計方法) を行う場合, 使用するカットオフ値を指定
			12.4	ランダム化後の複数の時点でアウトカム評価を実施する場合, 解析に使用する時点を明記
			12.5	複合アウトカムを使用する場合, 複合アウトカムの個々の構成要素をすべて定義
参加者のタイムライン	13	参加者の登録, 介入 (ランインとウォッシュアウトを含む), 評価, 来院のタイムスケジュール。概略図を強く推奨 (図参照)		
サンプルサイズ	14	研究目的を達成するために必要な推定参加者数とその決定方法 (サンプルサイズの計算を裏付ける臨床的および統計的仮定を含む)	14.1	治療群間の目標差 (例えば, 最小重要差) の定義と正当化
募集	15	目標サンプルサイズを達成するために十分な参加者登録を行うための手法		
方法: 介入の割付け (対照試験の場合) [Methods: assignment of interventions (for controlled trials)]				
割付け				
シーケンス生成	16a	割付けシーケンスの生成方法 (例: コンピュータで生成した乱数), および層別化のための因子のリスト。ランダム化シーケンスの予測可能性を減少させるため, 計画された制限 (例: ブロッキング) の詳細は, 参加者の登録または介入を割り当てる者が利用できない別の文書で提供されるべき		
割付け秘匿方法	16b	割付けシーケンスを実施する仕組み (例: 中央電話; 連番を付けた不透明な封をした封筒), 介入が割り当てられるまで順序を秘匿する手順の記述		
実施	16c	誰が割付けシーケンスを作成し, 誰が参加者を登録し, 誰が参加者を介入に割り付けるのか		
盲検化 (マスキング)	17a	介入への割付け後, 誰が盲検化されるのか (例: 試験参加者, 医療提供者, アウトカム評価者, データ解析者), またその方法		
	17b	盲検化されている場合, 盲検化を解除することが許される状況, および試験中に参加者に割り当てられた介入を明らかにする手順		

表 1-3

方法：データ収集, 管理, 分析 [Methods: data collection, management, and analysis]				
データ収集方法	18a	アウトカム, ベースライン, その他の試験データの評価と収集の計画 (データの質を向上させるための関連プロセス (例: 重複測定, 評価者のトレーニング), 評価尺度 (例えば, 質問票, 臨床検査) の説明, 信頼性と妥当性がわかっている場合はその説明も含む)。プロトコルに記載されていない場合, データ収集用紙の所在を言及	18a.1	調査標本に類似した集団における評価尺度の反応性について知られていることを記述
			18a.2	誰が主要アウトカムを評価するかを記述 (例: 看護師, 保護者)
	18b	介入を中止した, または介入プロトコルから逸脱した参加者のために収集される結果データのリストを含む, 参加者の定着と完全なフォローアップを促進する計画		
データ管理	19	データ入力, コード化, セキュリティ, 保管に関する計画 (データの質を高めるための関連プロセスを含む) (例: 二重データ入力, データ値の範囲チェック)。データ管理手順の詳細がプロトコルに記載されていない場合, どこに記載されているかを言及		
統計学的手法	20a	主要アウトカムおよび副次的アウトカムを分析するための統計学的方法。プロトコルにない場合, 統計解析計画の詳細がどこに記載されているかを言及	20a.1	主要アウトカムと副次的アウトカムの解析や解釈において, 多重性を考慮するための計画された方法があれば記述 (例: 共主要アウトカム, 同一アウトカムを複数の時点で評価する, アウトカムのサブグループ解析など)
	20b	追加解析の方法 (サブグループ解析や調整解析など)		
	20c	プロトコルの非遵守に関する解析集団の定義 (例: ランダム化解析), 欠損データを扱う統計学的手法 (例: 多重代入法)		
方法：モニタリング [Methods: monitoring]				
データモニタリング	21a	データモニタリング委員会 (DMC) の構成, 役割と報告体制の概要, スポンサーや競合利益から独立しているかどうかの記述, プロトコルに記載されていない場合は, 委員会の憲章に関する詳細な情報への言及。または, DMC が不要な理由の説明		
	21b	中間解析と試験中止のガイドラインの記述 (誰がこれらの中間解析結果にアクセスし, 試験中止の最終決定を下すかを言及)		
有害性	22	誘起された有害事象および自発的に報告された有害事象, ならびに試験介入または試験実施によるその他の予期せぬ影響を収集, 評価, 報告, 管理するための計画		
監査	23	試験実施状況を監査する頻度と手順 (もしあれば), そのプロセスが治験責任医師や治験依頼者から独立しているかどうか		
倫理と普及 [Ethics and dissemination]				
研究倫理承認	24	研究倫理委員会 / 施設審査委員会 (REC/IRB) の承認を求める計画		
プロトコルの修正	25	プロトコルの重要な変更 (例: 適格基準, アウトカム, 解析の変更) を関係者 (例: 試験責任医師, REC/IRB, 試験参加者, 試験登録, ジャーナル, 規制当局) に伝達する計画		

表 1-4

同意または承諾	26a	誰が試験参加候補者または承認された代理人からインフォームド・コンセントまたは同意を得るのか、またその方法（項目 32 を参照）		
	26b	該当する場合、補助的研究における参加者データおよび生物学的検体の収集および使用に関する追加の同意規定		
守秘義務	27	試験前、試験中、試験後の秘密を守るために、参加候補者および登録された参加者の個人情報がどのように収集、共有、保持されるか		
利益申告	28	試験全体および各試験実施施設の試験責任医師の金銭的利益およびその他の競合利益		
データへのアクセス	29	誰が最終的な試験データセットにアクセスできるかの記述、および試験責任医師のアクセスを制限する契約上の合意の開示		
補助的ケアと試験後のケア	30	該当する場合、補助的なケアや試験後のケア、試験参加によって損害を被った人への補償に関する規定		
普及方針	31a	試験責任医師および試験依頼者が、参加者、医療従事者、一般市民、その他関連団体に試験結果を伝達するための計画（例：出版、結果データベースへの報告、その他のデータ共有の取り決めなど）		
	31b	執筆資格のガイドラインおよび専門家による使用目的		
	31c	該当する場合、プロトコル全体、参加者レベルのデータセット、および統計コードへの一般公開を許可する計画		
倫理と普及 [Ethics and dissemination]				
インフォームド・コンセント資料	32	参加者および承認された代理人に提供される同意書およびその他の関連文書のモデル		
生物学的検体	33	該当する場合、今回の試験および将来の補助的研究における遺伝学的または分子学的分析のための生物学的検体の採取、検査室での評価および保管の計画		

略語：SPIRIT; Standard Protocol Items: Recommendations for Interventional Trials, TIDieR; Template for Intervention Description and Replication

a: このチェックリストは、SPIRIT 2013 声明のガイドライン 1, 2 と合わせて読むことを強く推奨する。SPIRIT 2013 声明チェックリストは、クリエイティブ・コモンズ・ライセンスの下で配布されている。

関連性を含む) について報告するよう推奨される。大まかな概念レベルでは、アウトカムのドメインとは、アウトカム（例：疼痛）を説明するために使用される名称や概念を指す^{13,38}。ドメインという言葉は、文献上ではコンストラクトやアトリビュートという言葉と密接に関連し、時には同等に使われることもある³⁹。アウトカムの完全な定義は試験プロトコルで提供されることが期待されるが（SPIRIT 2013 声明チェックリスト項目 12 で推奨されている²）、試験の主要アウトカムとしてアウトカムドメインを選択する根拠も、提案された試験の目的を支えるも

のであるため、伝えることが不可欠である。

根拠として重要な点には、

- (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、または医療費負担者）にとってのアウトカム領域の重要性、
- (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、かつ実行可能に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（すなわち、特定の臨床領域についてすべての試験で測定されるべき標準化されたアウト

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に対応; 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記述			
		• 記述的名称	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総死亡率
		• 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, および関連する場合, 比較される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 al³⁸⁾, Mayo-Wilson et al¹³⁾, Chan et al²⁾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。疼痛のように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³⁹⁾。

カムの合意されたセット)に由来するかどうかとも報告されるべきである⁴⁰⁻⁴⁴⁾。

項目 12.2. 主要アウトカムの分析指標が参加者内の変化を表す場合、個人における最小限の重要な変化を定義し、正当化する。

この項目は SPIRIT 2013 声明チェックリスト項目 12 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、前値（ベースラインで測定された値など）からの個人内変化を表す場合、関連する評価尺度について最小重要変化量 (MIC) とみなされるものの定義と正当性を示すべきである。SPIRIT Outcomes 2022 拡張版では、MIC は患者、臨床医、または関連する他者が重要と考える最小の患者内変化と定義された（一般的な代替用語は Box に示されている^{24,25,27)}。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⁾）を分析する予定の試験や n-of-1 デザインなど、参加者内の変化指標を使用するすべての試験で報告することが重要である⁴⁶⁾。

試験責任医師や試験関係者は、基準として MIC そのものに関心を持つかもしれないし、あるいは既知の MIC よりも大きな値に関心を持つかもしれない。計画された試験集団や設定に関して、評価尺度の MIC が不明な場合は、試験の一部として MIC を決定するための計画とともに報告されるべきである。試験計画書に選択された MIC の正当性を記載することは重要である。なぜなら、同じ評価尺度に対して、臨床的妥当性や方法論的質は、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が、多数の MIC が存在する可能性があるからである⁴⁷⁻⁵¹⁾。

項目 12.3. 収集したアウトカムデータが連続的であるが、カテゴリカル変数として分析される場合（集計方法）、使用されるカットオフ値を明記する。

この項目は、SPIRIT 2013 声明チェックリスト項目 12 を拡張したもので、使用するカットオフ値を前向きに記述し、連続尺度（または順序尺度）で収集されたアウトカムデータは、解析のためにカテゴリカル変数に変換すべきであることを著者に促すものである^{2,13,52)}。カットオフ値選択の根拠を説明することが推奨される。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⁵³⁾。試験プロトコル（および統計解析計画書）にカットオフ値を含めなかった場合¹⁶⁾、検出不能な多重検定（「p-ハッキング」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非報告を避けることができる^{7,13,54)}。

項目 12.4. ランダム化後、複数の時点でアウトカム評価を実施する場合は、解析に使用する評価時点を明記する。

この項目は、SPIRIT 2013 声明チェックリスト項目 12 を拡張したもので、ランダム化後の複数の時点でアウトカムの評価を実施する予定がある場合（例：ランダム化後 12 週間は毎日血圧を評価する試験）、主解析に使用する時点を明記するよう著者に促すものである。事前に評価時点を指定して解析することで、複数の評価時点をういた計画外の解析や、大きな結果や有意な結果が得られなかった時点を選択的に報告しない可能性を抑えることができる^{2,38)}。時点選択の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、臨床試験が情報提供するためにデザインされた管理上の決定に対して適切であるべきである⁵⁵⁾。

項目 12.5. 複合アウトカムを使用する場合は、複合アウトカムの個々の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から構成される。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^{20,21)}。複合アウトカムを使用する場合、試験プロトコルにおいて⁵³⁾、個々の構成アウトカムの完全な定義（SPIRIT 2013 声明チェックリスト項目 12 および SPIRIT-Outcomes 2022 拡張版チェックリスト項目 12.1~12.4）、およびこれらのアウトカムの組み合わせ方（例：どのような分析ステップで複合アウトカムの発生を定義するのか）の説明を含め、事前に特定し、正当化し、完全に定義する必要がある。複合アウトカムの使用と解釈は複雑で、文献で議論されており、重要な試験デザイン上の考慮事項である^{21,56)}。

項目 14.1. 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、SPIRIT 2013 声明チェックリスト項目 14 のサンプルサイズの計算を拡張したもので、解析の特定の時点における先験的な目標群間治療差を明記するよう著者に促すものである。目標群間差は、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とみなされる、解析の特定の時点における介入群間の主要アウトカムで検出される差として、サンプルサイズの計算に使用される値である³⁶⁾。The Difference Elicitation in Trials プロジェクトは、試験の目標差の選択、サンプルサイズの計算、報告について、エビデンスに基づいた広範なガイダンスを発表している^{36,37)}。目標とする差は、最小重要変化量（患者、臨床医、または関連する他者が重要であると認識する患者間の最小の差）^{24,28,29)}、または最小有益効果（患者が決定する介入の費用、有害性、不便さを正当化する最小の有益な効果）とすることができる³⁴⁾。目標差を報告することは、

- (1) 助成金の査読者が資金配分を正当化するのに十分な妥当性と臨床効果があるかどうかを判断するため、
- (2) 倫理委員会、規制当局、その他の査読者が、試験を承認するのに十分な有益性と有害性の比率があるかどうか、臨床的に重要な差を検出するのに十分な検出力があるかどうかを評価するために、試験プロトコルの段階で不可欠である。選択された目標差を決定する実用的または臨床的な要因は様々であり（例：主要アウトカムを評価するために使用される評価尺度について、信頼できる最小重要差が利用可能であること）⁴⁷⁾、利用可能な選択肢は多数あるため（例：複数の最小重要差のうちの 1 つ、またはパイロットスタディに基づく値）⁴⁷⁾、選択された目標差が選択された理由を説明することが重要である^{26,48,49)}。

データ収集、管理、分析の記述に関する SPIRIT-Outcomes 2022 拡張版チェックリスト項目【SPIRIT-Outcomes 2022 Extension Checklist Items for the Descriptions of the Data Collection, Management, and Analysis】

項目 18a.1. 調査標本に類似した集団における評価尺度の反応性について知られていることを記述する。

この項目は、SPIRIT 2013 声明チェックリスト項目 18 で、評価尺度（例：質問票、臨床検査）の信頼性と妥当性の記述を求めたものを拡張し、研究手段の応答性についても報告するものである。応答性とは、評価されるアウトカムの経時的変化を正確に検出し測定する評価尺度（すなわち、経時的変化や典型的な治療効果を測定する尺度）⁵⁷⁾ の能力を指す^{25,32,33)}。応答性は、診断精度が重要である識別的な評価尺度（例：診断テスト）にはあまり関係しない。測定特性（応答性など）が標本と類似した集団（または少なくとも標本と実質的に異なる集団）で評価されているかどうかを記述することは有用である。なぜなら、評価尺度の測定特性は文脈特異的であり、エビデンスがない限り異なる集団間（例：異なる健康問題を持つ集団や異なる年齢層）で一般化が可能だとは仮定できないからである⁵⁸⁾。研究集団に対する測定特性が不明な場合（例：青年を対象とした試験で成人用に開発された尺度を使用する場合）、この尺度が他の利用可能な選択肢よりも有用であり、望ましいと予想される理由と、他に利用可能な尺度があるかどうかについての根拠とともに記載する。

項目 18a.2. 主要アウトカムを誰が評価するかを記述する（例：看護師、保護者）。

この項目は SPIRIT 2013 声明チェックリスト項目 18 を拡張したもので、アウトカムデータの評価と収集の計画の説明の一部として、誰が主要アウトカムを評価するかを記述する。患者の治療割り付けと新たな試験結果に対するアウトカム評価者の盲検化は、SPIRIT 2013 声明チェックリスト項目 17a でカバーされている。多くの種類のアウトカム（例：行動学的アウトカム、心理学的アウトカム）において、誰がアウトカムを評価するかによって、実質的に異なる回答、ひいては異なる試験結果が得られる可能性がある。このようならばつきは、評価者の訓練や経験の違い、視点の違い、患者の想起などに起因する可能性がある^{59,61)}。臨床医、患者、非臨床医の観察者から報告された臨床アウトカムの評価、または実績に基づく評価は、米国食品医薬品局によって、臨床医報告、患者報告、観察者報告、実績アウトカ

ムに分類されている⁶²⁾。様々な人によって評価される可能性のあるアウトカムについては、試験の文脈で行われたアウトカム評価者の選択に関する説明を提供すべきである。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値）については、この情報はあまり意味がない。

項目 20a.1. 主要アウトカムと副次的アウトカムの解析または解釈における多重性考慮のための計画された方法を記述する（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、またはアウトカムのサブグループ解析）

この項目は SPIRIT 2013 声明チェックリスト項目 20a を拡張したもので、著者にアウトカムの解析や解釈に関連する多重性を考慮するために計画された方法を記述するよう促すものである。アウトカムの多重性の問題は試験において一般的であり、特に、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り複数の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）、解析の対象となる副次的アウトカムが多数存在する場合には注意が必要である⁶³⁾。このような多重性を考慮するために使用される計画された方法には、統計学的方法（例：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）が含まれる^{53,64)}。このような情報は、プロトコルに簡潔に記載されることもあれば、統計解析計画書に詳細に記載されることもある¹⁶⁾。

考察【Discussion】

SPIRIT-Outcomes 2022 拡張版は、試験プロトコルにおけるアウトカム特異的情報報告のためのエビデンスとコンセンサスに基づくガイダンスを提供するもので、SPIRIT 2013 声明チェックリストに9つの報告項目を追加し、報告に関する推奨事項を CONSORT-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つの拡張ガイドラインを整合させることで、試験プロトコルから完成した試験報告書までの一貫した連続的な報告が作成され、研究者による試験プロトコルと試験報告書

の作成が容易になり、また重要なこととして、研究者による試験プロトコルの遵守状況の評価が可能になる。

SPIRIT 2013 声明と同様に^{1,2)}、SPIRIT-Outcomes 2022 拡張版は、試験デザインや集団に関係なく、試験プロトコルの内容に適用される。現行の推奨事項も同様に、どのように情報を記載すべきかについて規定するものではなく、地域要件、試験プロトコルが意図する目的と対象者（例：規制当局への提出と学術誌への発表）、関連する試験文書（例：統計解析計画書）の公開状況によって異なる¹⁶⁾。

他の文書が利用可能な場合は、試験プロトコルにチェックリスト項目の構成要素を簡潔に記述または概説し、統計解析計画書などの別の文書を参照することで、それらの存在、所在、利用しやすさを周知させる。SPIRIT-Outcomes 2022 拡張版の有用性と効果を最大化するために、CONSORT-Outcomes 2022 拡張版と併用することができる¹⁷⁾。さらに、試験登録（例：ClinicalTrials.gov）に SPIRIT-Outcomes 2022 拡張版チェックリスト項目を含めることで、実施率が高まる可能性がある。出版社や編集者が著者のために SPIRIT-Outcomes 2022 拡張版を支持し、実施することも、導入が増加する可能性がある。試験プロトコルに記載された情報は、試験登録や統計解析計画で報告された情報と一致すべきである。試験プロトコルは、自由かつ前向きにアクセス可能であるべきである（例：出版物または恒久的なオンラインリポジトリを介して）。

SPIRIT-Outcomes 2022 拡張版の利用者は、これらの追加チェックリスト項目は、アウトカム特異的報告項目として最低限必要なものであり、試験の透明性、再現性、批判的評価を促進し、結果の選択的非報告を制限するために、SPIRIT 2013 声明ガイドラインに追加されていることに留意すべきである（付録の eTable 5 ※）。場合によっては、試験プロトコルに患者報告アウトカムを記載する際のガイダンスとして SPIRIT-PRO 拡張版⁵⁾ を参照することも重要である。SPIRIT-Outcomes 2022 拡張版チェックリスト項目の中には、主要アウトカムについてのみ必須の報告項目としてコンセンサスを得たものがあるが、他の試験のアウトカムについても報告することが重要である場合がある。SPIRIT-Outcomes 2022 拡張版に従う著者は、チェックリストに記入する際

に、なぜその項目が自分の試験に関連しないのかを明記すべきである。例えば、系統的に評価されたアウトカムを報告するためのこの拡張版チェックリストは、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、または事前に規定されていないアウトカムには適用できないことがある。

SPIRIT-Outcomes 2022 拡張版の主な利用者は、試験プロトコル作成者、倫理審査委員会、雑誌編集者であると予想される。これらのグループが SPIRIT-Outcomes 2022 拡張版を使用することで、試験の有用性と透明性が向上し、CONSORT-Outcomes 2022 拡張版と併用することで、試験報告における結果の選択的非報告のリスクを低減できる可能性がある¹⁷⁾。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。SPIRIT-Outcomes 2022 拡張版は、EQUATOR ネットワークや SPIRIT のウェブサイトを含め、以前に概説したように普及される予定である¹⁴⁾。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出することができ⁶⁵⁾、今後のアップデートに反映される。

研究限界 [Limitations]

本研究にはいくつかの限界がある。

第一に、含まれるチェックリストの項目は、ほとんどの潜在的利益といくつかの有害性を含む、体系的に収集されたアウトカムに適しているが、他の項目は体系的に評価されていない有害性の報告に適用できる可能性がある⁶⁶⁾。

第二に、これらのチェックリスト項目はまだ

SPIRIT の主要なチェックリストに統合されていないため、著者や編集者によっては、複数のチェックリストを探し、使用することが負担になり、実施に影響が出る可能性がある⁶⁷⁾。これらの追加項目を SPIRIT の主要なチェックリストに統合する今後の取り組みによって、実際の実施が促進されるかもしれない。

第三に、より広範な研究コミュニティにおけるユーザビリティを向上させる目的で、大規模で多様な国際的な専門家およびエンドユーザーグループがこれらの推奨事項の策定に関与したが、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サインアップした）、デルファイ調査の結果は非回答バイアスの影響を受けた可能性がある。

第四に、コンセンサス会議のパネリストは、ランダム化臨床試験の実施、監督、報告、試験結果の使用に関連する専門知識と役割に基づいて、意図的に抽出された。コンセンサス会議のパネリストに代表されない人々（例：北米やヨーロッパ以外の試験実施者）の意見は異なる可能性がある。アウトカム報告ガイダンスの厳密なスコーピングレビュー^{6, 18)}を含む、本ガイドラインの作成に用いられた体系的でエビデンスに基づいたアプローチ^{14, 15)}は、これらの制限の潜在的な影響を緩和するのに役立つであろう。

結論 [Conclusions]

この SPIRIT-Outcomes 2022 は、SPIRIT 2013 声明を拡張したもので、すべての試験プロトコルで扱うべきアウトカムに特異的な9項目を示しており、試験の有用性、再現性、透明性を高め、試験結果の選択的非報告のリスクを最小限に抑える可能性がある。

参考文献 [REFERENCES]

1. 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
2. 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
3. SPIRIT. SPIRIT statement website. Accessed April 13, 2021. <https://www.spirit-statement.org/>
4. International Conference on Harmonisation. E6: good clinical practice. Accessed November 22, 2022. <https://www.ich.org/page/efficacy-guidelines>
5. 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
6. 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
7. 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
8. Yordanov Y, Dechartres A, Atal I, *et al.*: Avoidable waste of research related to outcome planning and reporting in clinical trials. *BMC Med.* 2018; **16**(1): 87. doi:10.1186/s12916-018-1083-x
9. 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
10. 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
11. 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
12. 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
13. 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
14. 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
15. 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
16. 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
17. Butcher NJ, Monsour A, Mew EJ, *et al.*: Guidelines for reporting outcomes in trial reports: the CONSORT-Outcomes 2022 extension. *JAMA.* 2022;**328**(22): 2252-2264. doi:10.1001/jama.2022.21022
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. Ferreira-González I, Permyanier-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
21. 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
22. 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
23. 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
24. Mokkink LB, Terwee CB, Patrick DL, *et al.*: COSMIN checklist manual. Accessed February 28, 2022. https://fac.ksu.edu.sa/sites/default/files/cosmin_checklist_manual_v9.pdf
25. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine.* Cambridge University Press; 2011. doi:10.1017/CBO9780511996214
26. 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
27. 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
28. de Vet HC, Beckerman H, Terwee CB, Terluin B, Bouter LM. Definition of clinical differences. *J Rheumatol.* 2006; **33**(2): 434-435.

29. 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
30. 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/>
31. 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
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. 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
34. 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
35. 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
36. 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
37. 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
38. 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
39. OMERACT. OMERACT domain resources. Accessed May 19, 2022. <https://omeract.org/domain-selection/>
40. 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
41. Gargon E, Gorst SL, Williamson PR. Choosing important health outcomes for comparative effectiveness research: 5th annual update to a systematic review of core outcome sets for research. *PLoS One.* 2019; **14**(12): e0225980. doi:10.1371/journal.pone.0225980
42. Williamson PR, Altman DG, Bagley H, et al.: The COMET Handbook: version 1.0. *Trials.* 2017; **18** (suppl 3): 280. doi:10.1186/s13063-017-1978-4
43. Kirkham JJ, Davis K, Altman DG, et al.: Core Outcome Set-STAndards for Development: the COS-STAD recommendations. *PLoS Med.* 2017; **14**(11): e1002447. doi:10.1371/journal.pmed.1002447
44. Gorst SL, Prinsen CAC, Salcher-Konrad M, Matvienko-Sikar K, Williamson PR, Terwee CB. Methods used in the selection of instruments for outcomes included in core outcome sets have improved since the publication of the COSMIN/COMET guideline. *J Clin Epidemiol.* 2020; **125**: 64-75. doi:10.1016/j.jclinepi.2020.05.021
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. Devji T, Carrasco-Labra A, Qasim A, et al.: Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study. *BMJ.* 2020; **369**: m1714. doi:10.1136/bmj.m1714
52. Mayo-Wilson E, Fusco N, Li T, Hong H, Canner JK, Dickersin K; MUDS investigators. Harms are assessed inconsistently and reported inadequately part 1: systematic adverse events. *J Clin Epidemiol.* 2019; **113**: 20-27. doi:10.1016/j.jclinepi.2019.04.022
53. International Conference on Harmonisation. Harmonised Tripartite Guideline: statistical principles for clinical trials E9: current step 4 version. Accessed November 19, 2022. https://database.ich.org/sites/default/files/E9_Guideline.pdf
54. 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
55. 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/>
56. Freemantle N, Calvert M. Weighing the pros and cons for composite outcomes in clinical trials. *J Clin Epidemiol.* 2007; **60**(7): 658-

659. doi:10.1016/j.jclinepi.2006.10.024
57. Hankins M. How discriminating are discriminative instruments? *Health Qual Life Outcomes*. 2008; **6**(1): 36. doi:10.1186/14777525-6-36
58. 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
59. 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
60. Cohen SB, Strand V, Aguilar D, Ofman JJ. Patient versus physician-reported outcomes in rheumatoid arthritis patients treated with recombinant interleukin-1 receptor antagonist (anakinra) therapy. *Rheumatology (Oxford)*. 2004; **43**(6): 704-711. doi:10.1093/rheumatology/keh152
61. 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
62. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) resource. Accessed May 19, 2022. <https://www.ncbi.nlm.nih.gov/books/NBK338448>
63. 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
64. 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
65. REDCap Research. SPIRIT-Outcomes feedback form. Accessed February 28, 2022. <https://redcapexternal.research.sickkids.ca/surveys/?s=7N847LCEJF>
66. 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
67. 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