

シリーズ EQUATOR Network が提供するガイドラインの紹介

Reporting of surrogate endpoints in randomised controlled trial reports (CONSORT-Surrogate): extension checklist with explanation and elaboration の和訳

著者 Anthony Muchai Manyara^{1,2}, Philippa Davies³, Derek Stewart⁴, Christopher J Weir⁵, Amber E Young³, Jane Blazeby^{3,6,7}, Nancy J Butcher^{8,9}, Sylwia Bujkiewicz¹⁰, An-Wen Chan^{11,12}, Dalia Dawoud^{13,14}, Martin Offringa^{8,15}, Mario Ouwens¹⁶, Asbjørn Hróbjartsson^{17,18}, Alain Amstutz^{19,20,21}, Luca Bertolaccini²², Vito Domenico Bruno²³, Declan Devane^{24,25}, Christina D C M Faria²⁶, Peter B Gilbert²⁷, Ray Harris⁴, Marissa Lassere²⁸, Lucio Marinelli^{29,30}, Sarah Markham^{4,31}, John H Powers 3rd³², Yousef Rezaei^{33,34,35}, Laura Richert³⁶, Falk Schwendicke³⁷, Larisa G Tereshchenko³⁸, Achilles Thoma³⁹, Alparslan Turan⁴⁰, Andrew Worrall⁴, Robin Christensen⁴¹, Gary S Collins⁴², Joseph S Ross^{43,44}, Rod S Taylor^{1,45}, Oriana Ciani⁴⁶

翻訳 馬場 亜沙美 (BABA Asami)^{1*}, 鈴木 直子 (SUZUKI Naoko)¹, 田中 瑞穂 (TANAKA Mizuho)¹, 山本 和雄 (YAMAMOTO Kazuo)¹

Key Words: ヒト臨床試験, ランダム化比較試験, 特定保健用食品, 機能性表示食品, CONSORT 拡張版チェックリスト, 代替評価項目, SPIRIT

Japanese translation of “Reporting of surrogate endpoints in randomised controlled trial reports (CONSORT-Surrogate): extension checklist with explanation and elaboration”

Authors:

Anthony Muchai Manyara^{1,2}, Philippa Davies³, Derek Stewart⁴, Christopher J Weir⁵, Amber E Young³, Jane Blazeby^{3,6,7}, Nancy J Butcher^{8,9}, Sylwia Bujkiewicz¹⁰, An-Wen Chan^{11,12}, Dalia Dawoud^{13,14}, Martin Offringa^{8,15}, Mario Ouwens¹⁶, Asbjørn Hróbjartsson^{17,18}, Alain Amstutz^{19,20,21}, Luca Bertolaccini²², Vito Domenico Bruno²³, Declan Devane^{24,25}, Christina D C M Faria²⁶, Peter B Gilbert²⁷, Ray Harris⁴, Marissa Lassere²⁸, Lucio Marinelli^{29,30}, Sarah Markham^{4,31}, John H Powers 3rd³², Yousef Rezaei^{33,34,35}, Laura Richert³⁶, Falk Schwendicke³⁷, Larisa G Tereshchenko³⁸, Achilles Thoma³⁹, Alparslan Turan⁴⁰, Andrew Worrall⁴, Robin Christensen⁴¹, Gary S Collins⁴², Joseph S Ross^{43,44}, Rod S Taylor^{1,45}, Oriana Ciani⁴⁶

Translators:

Asami Baba^{1*}, Naoko Suzuki¹, Mizuho Tanaka¹, Kazuo Yamamoto¹

*Correspondence author: Asami Baba

Keywords: Human clinical trials, Randomized controlled trials, Foods for Specified Health Uses (FOSHU), Foods with Functional Claims (FFC), CONSORT extension checklist, Surrogate endpoints, SPIRIT

Affiliations (Authors):

¹ MRC/CSO Social and Public Health Sciences Unit, School of Health and Wellbeing, University of Glasgow, Glasgow, UK

² Global Health and Ageing Research Unit, Bristol Medical School, University of Bristol, Bristol, UK

³ Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

⁴ Patient author, UK

* 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

- ⁵ Edinburgh Clinical Trials Unit, Usher Institute, University of Edinburgh, Edinburgh, UK
- ⁶ Bristol NIHR Biomedical Research Centre, Bristol, UK
- ⁷ University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, UK
- ⁸ Child Health Evaluative Sciences, Hospital for Sick Children Research Institute, Toronto, ON, Canada
- ⁹ Department of Psychiatry, University of Toronto, Toronto, ON, Canada
- ¹⁰ Biostatistics Research Group, Department of Population Health Sciences, University of Leicester, Leicester, UK
- ¹¹ Women's College Research Institute, Toronto, ON, Canada
- ¹² Department of Medicine, University of Toronto, Toronto, ON, Canada
- ¹³ Science, Evidence, and Analytics Directorate, Science Policy and Research Programme, National Institute for Health and Care Excellence, London, UK
- ¹⁴ Faculty of Pharmacy, Cairo University, Cairo, Egypt
- ¹⁵ Department of Paediatrics, University of Toronto, Toronto, ON, Canada
- ¹⁶ AstraZeneca, Mölndal, Sweden
- ¹⁷ Centre for Evidence-Based Medicine Odense and Cochrane Denmark, Department of Clinical Research, University of Southern Denmark, Odense, Denmark
- ¹⁸ Open Patient data Explorative Network, Odense University hospital, Odense, Denmark
- ¹⁹ CLEAR Methods Centre, Division of Clinical Epidemiology, Department of Clinical Research, University Hospital Basel and University of Basel, Basel, Switzerland
- ²⁰ Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway
- ²¹ Bristol Medical School, University of Bristol, Bristol, UK
- ²² Department of Thoracic Surgery, IEO, European Institute of Oncology IRCCS, Milan, Italy
- ²³ IRCCS Galeazzi-Sant'Ambrogio Hospital, Department of Minimally Invasive Cardiac Surgery, Milan, Italy
- ²⁴ University of Galway, Galway, Ireland
- ²⁵ Health Research Board-Trials Methodology Research Network, University of Galway, Galway, Ireland
- ²⁶ Department of Physical Therapy, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
- ²⁷ Fred Hutchinson Cancer Centre, Seattle, WA, USA
- ²⁸ St George Hospital and School of Population Health, University of New South Wales, Sydney, NSW, Australia
- ²⁹ Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genova, Genoa, Italy
- ³⁰ IRCCS Ospedale Policlinico San Martino, Genoa, Italy
- ³¹ Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
- ³² George Washington University School of Medicine, Washington, DC, USA
- ³³ Heart Valve Disease Research Centre, Rajaie Cardiovascular Medical and Research Centre, Iran University of Medical Sciences, Tehran, Iran
- ³⁴ Ardabil University of Medical Sciences, Ardabil, Iran
- ³⁵ Behyan Clinic, Pardis New Town, Tehran, Iran
- ³⁶ University of Bordeaux, Centre d'Investigation Clinique-Epidémiologie Clinique 1401, Research in Clinical Epidemiology and in Public Health and European Clinical Trials Platform & Development/French Clinical Research Infrastructure Network, Institut National de la Santé et de la Recherche Médicale/Institut Bergonié/Centre Hospitalier Universitaire Bordeaux, Bordeaux, France
- ³⁷ Charité Universitätsmedizin Berlin, Berlin, Germany
- ³⁸ Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, USA
- ³⁹ McMaster University, Hamilton, ON, Canada
- ⁴⁰ Department of Outcomes Research, Anaesthesiology Institute, Cleveland Clinic, OH, USA
- ⁴¹ Section for Biostatistics and Evidence-Based Research, the Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen and Research Unit of Rheumatology, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, Odense, Denmark
- ⁴² UK EQUATOR Centre, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Oxford, UK
- ⁴³ Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA
- ⁴⁴ Section of General Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA
- ⁴⁵ Robertson Centre for Biostatistics, School of Health and Well Being, University of Glasgow, Glasgow, UK
- ⁴⁶ Centre for Research on Health and Social Care Management, Bocconi University, Milan 20136, Italy

Affiliated institution (Translators)

¹ ORTHOMEDICO Inc.

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

本稿について

本稿は、EQUATOR Network が提供するガイドラインの一つである「Reporting of surrogate endpoints in randomised controlled trial reports (CONSORT-Surrogate): extension checklist with explanation and elaboration」を翻訳・要約したものである。なお、付録および補足資料は原文 (<https://www.bmj.com/content/386/bmj-2023-078524>) にアクセスして参照する。

要約ポイント [Summary points]

- ランダム化比較試験では、特に医薬品や生物学的製剤の規制承認や医療技術評価の場面において、最終的な臨床アウトカムの代替として代替エンドポイントが用いられることが多い。
- しかし、試験における代替エンドポイントの使用は、介入の有効性や目標とする臨床転帰に対する効果の解釈に関して誤解を招く可能性があり、さらに安全性、すなわち有害事象に関する情報が十分に得られない場合がある。
- 本稿では、代替エンドポイントを主要アウトカムとして用いた試験報告の質を向上させ、ひいてはより適切な患者ケア、医療上の意思決定および政策決定に資することを目的として策定された拡張ガイドラインである CONSORT-Surrogate について概説する。
- 臨床試験の著者、ジャーナル編集者および査読者は、試験報告の方法論的完全性、透明性、再現性、解釈の妥当性ならびに結果の有用性を高めるために、CONSORT-Surrogate を活用することが推奨される。

要旨 [Abstract]

ランダム化比較試験では、効率性向上（試験期間の短縮、標本サイズの縮小、ひいては研究コストの削減）や倫理的・実務的理由から、目標アウトカム（試験参加者、臨床医、その他の利害関係者にとって直接的な関心および関連性を有する結果、例：全死因死亡率）に代わる代替エンドポイントが一般的に用いられている。しかし、代替エンドポイントへの依存は、介入の治療効果に関する不確実性を高め、さらに介入の有害性に関する十分な情報を提供できない可能性があることから、代替エンドポイントを用いた試験報告の改善を求める声が高まっている。本報告は、代替エンドポイントを主要アウトカムとして用いる試験のための合意に基づく報告ガイドライン、すなわち Consolidated Standards of Reporting Trials (CONSORT) 拡張チェックリスト：CONSORT-Surrogate を提示するものである。本拡張版は、CONSORT 2010 チェックリストの9項目を修正するとともに、新たに2項目を追加した構成となっている。各項目については、具体例および解説を付している。全ての関係者（試験実施者・スポンサー、ジャーナル編集者・査読者、研究倫理審査委員、資金提供者を含む）に対し、代替エンドポイントを用いた試験報告において本拡張版を活用することを推奨する。本チェックリストの使用は、試験結果の透明性、解釈可能性および有用性の向上に資するとともに、最終的には研究資源の不必要な消費、すなわち研究の無駄の削減につながることを期待される。

序論 [Introduction]

介入の有効性または効果を判断するには、対象となる関心のある転帰（例：全死因死亡率）に対する介入の効果を評価した、適切に設計・実施・報告されたランダム化比較試験（本稿では「試験」と呼称する）からのエビデンスが必要である¹⁾。試験の不十分な報告は意思決定における有用性を低下させ、研究の無駄を増大させる問題につながる^{2,3)}。報告ガイドラインの使用は、試験エビデンスの有用性向

上および研究の無駄の削減に有効であることが示されている³⁾。

Consolidated Standards of Reporting Trials (CONSORT) 声明は、並行群間試験の報告に広く用いられている25項目から成るチェックリストである⁴⁾。CONSORT チェックリストは試験報告の完全性を向上させたが⁵⁾、あらゆる種類の試験に十分対応できるものではない。このため、CONSORT 拡張版（項目を修正または追加したチェックリスト）が開発

表1 ランダム化比較試験における代替エンドポイントの例

項目	例 1 ¹⁰⁾	例 2 ¹¹⁾	例 3 ¹²⁾
代替エンドポイントの領域	血圧	腫瘍反応	BMI
測定変数または特定測定	昼間活動時の収縮期血圧	固形腫瘍の反応評価基準 (RECIST) バージョン 1.1 に基づき、造影剤強化 CT または MRI を用いた独立中央審査により評価	電子健康記録から取得した体重と身長から算出
特定指標	ベースラインからの変化量	時点における測定値	ベースラインからの変化量
集計方法	連続的アウトカム: 平均値	二値アウトカム: 頻度 (%)	年齢・性別別 BMI 中央値からのパーセンテージ
時点	ランダム化後 2 か月	中央値 15.3 か月	親の同意取得または研究登録後 2 年
著者による目標アウトカム	脳卒中, 冠状動脈性心臓病, 心不全, 全死因死亡率	全生存率	糖尿病, 肝臓病, 喘息, 心臓病, がん, 健康関連の生活の質の低下, 行動障害, 心理社会的機能障害

最初の 5 行は, Butcher ら⁷⁾, Chan ら¹³⁾, Mayo-Wilson ら¹⁴⁾, および Zarin ら¹⁵⁾ を参考にした, 定義されたアウトカムのコア要素を表している。

されている (例: CONSORT-PRO (患者報告アウトカム)⁶⁾, CONSORT-Outcomes⁷⁾)。しかし, 既存の拡張版はいずれも, 代替エンドポイントを用いる試験に対する具体的な指針を提供していない。

代替エンドポイントは, 試験において目標アウトカムの代替として機能する^{8,9)}。表 1 に, 試験で適用可能な代替エンドポイントの例をいくつか示す。

代替エンドポイントは, 試験の効率化 (例: 追跡期間の短縮, サンプルサイズの縮小, ひいては試験全体のコスト削減)をはじめ, 実現可能性, 実用性, 倫理的および科学的理由などから頻繁に用いられる¹⁶⁾。疾患や医療分野, 代替エンドポイントの定義にもよるが, 試験の 20~78% が主要アウトカムとして代替エンドポイントを採用していると推定されている¹⁷⁻²⁰⁾。

しかしながら, 目標アウトカムに関するデータが不足している状況では, 試験における代替エンドポイントの使用には議論の余地があり, 臨床的および方針的意思決定において根本的な限界を伴う。具体的には, 介入が目標アウトカム (ならびに臨床的有効性・効果, 費用対効果) に及ぼす真の効果に関する不確実性の増大, ならびに介入の有害性に関する十分な情報が提供されない可能性が挙げられる。これは, 代替エンドポイントを用いた試験が通常, より小規模なサンプルサイズおよび短い追跡期間で実施されるためである¹⁶⁾。

したがって, 代替エンドポイントに依存する試験の報告改善が求められており, 代替エンドポイント使用の明示的な表明とその根拠, ならびに潜在的な限界に対する十分な考慮が必要とされている²⁰⁻²³⁾。代替エンドポイントを用いた試験報告における継続的な不備を踏まえ, SPIRIT/CONSORT-Surrogate プロジェクトが設立され, 代替エンドポイントを主要アウトカムとする試験向けに SPIRIT および CONSORT の拡張が開発された。SPIRIT の拡張版である SPIRIT-Surrogate は Manyara らによって提示されている²⁴⁾。本稿では, CONSORT の拡張版である CONSORT-Surrogate チェックリストおよびその詳細説明文書を提示する。表 2 は, この拡張版で使用される用語集を示す。

CONSORT-Surrogate の適用範囲および使用法

[Scope and use of CONSORT-Surrogate]

ボックス 1 では CONSORT-Surrogate の適用範囲と使用法を要約する。本拡張版は, あらゆる定義に基づく代替エンドポイントが主要アウトカムとして用いられる全ての試験タイプおよびフェーズ (複合アウトカムの一部として代替エンドポイントが使用される場合を含む) の報告に適用すべきである。主要アウトカムが介入の評価および試験の結論を左右することを踏まえ, 本拡張版はこの側面に焦点を当てている。

表2 CONSORT-Surrogate で使用される用語集

用語	定義
複合アウトカム	2つ以上のコンポーネントアウトカムから成るアウトカム(例:死亡または非致死性脳卒中を発症した参加者の割合)。コンポーネント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 ²⁵⁾ 。
CONSORT	試験報告の統合基準。完了したランダム化比較試験の報告チェックリスト。
CONSORT-Surrogate	代替エンドポイントを主要アウトカムとして用いる試験を報告するために用いられる、拡張版CONSORTチェックリスト。
主要アウトカム	介入の有効性を評価する上で試験チームが最も重要かつ実現可能と考える事前定義されたアウトカム。標本サイズ計算や試験結論の根拠となる。主要エンドポイントとも呼ばれる ^{7,26)} 。
副次的アウトカム	追加的な介入効果を測定するために試験プロトコルで事前指定されたアウトカム ⁶⁾ 。研究デザインに基づき、確認的または探索的(すなわち、偽陽性結果増加に対する統計的調整)となる場合がある。
SPIRIT	介入試験のための標準プロトコル項目:推奨事項。完了したランダム化比較試験プロトコルの報告チェックリスト。
SPIRIT-Surrogate	代替エンドポイントを主要アウトカムとして使用する試験プロトコルの報告に用いられる拡張版SPIRITチェックリスト。
代替エンドポイント ^{†)}	患者の自覚症状、機能状態、生存期間を直接測定する代わりに試験で使用されるエンドポイント。目的である実際の臨床的ベネフィットを測定するものではないが、疫学的、治療学的、病態生理学的、その他の科学的根拠に基づき、臨床的ベネフィットまたは有害性に対する治療効果を予測すると期待されるもの ⁸⁾ 。
目標アウトカム	試験参加者、患者、臨床医、試験実施者、その他の利害関係者にとって直接的な関心と関連性を持つアウトカム ²⁷⁾ 。

*SPIRIT/CONSORT-Surrogate 拡張プロジェクト研究(e-Delphi 法および e-調査を含む)では、試験参加者、臨床医、試験実施者、規制当局、支払者といった異なる利害関係者間における代替エンドポイントの定義を調査した。本研究の結果および定義に関する考察の詳細は別報で報告されている²⁷⁾。

†「代替」と共に用いられる他の記述用語には、「アウトカム」「マーカー」「測定値」「観察値」「パラメータ」がある。文献では「早期」「代替」「プロキシ」「代替エンドポイント」「アウトカム」「測定値」「マーカー」とも呼ばれることがある。

本拡張版は報告すべき最小限の推奨項目を提供するものであるが、試験著者は代替エンドポイント試験の透明性向上や結果解釈に資する追加情報を提供してもよい。重要な点として、本拡張は推奨項目に適合させるために試験チームの設計や計画の変更を義務付けるものではない。すなわち、試験著者は実施された、または計画された内容を明示するだけでよいが、可能な限り全項目の実装を検討することが強く推奨される。

ボックス1では、本拡張の適用範囲および使用方法に関する追加的側面を示す。付録表A1は、本拡張項目を裏付ける試験報告における代替エンドポイントの設計および報告に関する主要な方法論的考慮事項を提示する(Web appendix 1)。

ボックス 1
CONSORT-Surrogate の適用範囲と使用方法
使用適格性
 代替エンドポイント(定義を問わず)を主要アウ

トカムとして用いる全ての介入ランダム化比較試験。代替エンドポイントが主要複合アウトカムの一部である場合も含む。

最低要件

本拡張版は報告すべき最小限の項目セットであるが、試験著者は透明性・明確性・結果解釈の向上のため追加情報を提供できる。

代替エンドポイントの検証手法は対象外

代替エンドポイントの検証手法や使用・引用すべき指標の評価は、本拡張の適用範囲外である。

目標アウトカム

試験チームは、目標アウトカム(副次的アウトカムとして)の収集と介入効果の報告を検討すべきである。この情報は、その後の代替エンドポイントの検証解析や介入による潜在的な有害事象の評価を支援し得る。

報告項目の順序の柔軟性

項目は統合可能であり、本拡張で提案された項目とは異なるセクションで報告してもよい。特定の項目セクションは必須要件ではなく推奨事項である。

拡張項目の外挿

本拡張は試験向けに開発されたが、非ランダム化試験、観察研究、および代替エンドポイントを用いるその他の研究の報告にも関連し得る。

CONSORT = Consolidated Standards of Reporting Trials (試験報告の統合基準)

CONSORT-Surrogate の開発

[Development of CONSORT-Surrogate extension]

SPIRIT-Surrogate と並行して実施された CONSORT-Surrogate の開発は、健康研究報告ガイドライン策定のための Enhancing the QUality and Transparency Of health Research (EQUATOR: 健康研究の質と透明性の向上) ネットワークのガイダンスに基づき、以下の 4 段階を経て進められた²⁸⁾。この開発は EQUATOR ネットワークのウェブサイト²⁹⁾ で事前登録され、プロトコルが公開された³⁰⁾。

フェーズ 1 では、文献レビューを実施し、既存文献から代替エンドポイントを用いた試験の報告項目を統合するとともに、代替エンドポイントに関する内容専門家を特定した (探索的レビュー)。また、代替エンドポイントを主要アウトカムとして用いた最近の試験の試験責任者を特定し、e-Delphi 調査への招待対象とした (対象限定レビュー)。文献レビューのプロトコルは別報で公表済みである³¹⁾。

探索的レビューの検索は 2022 年 3 月から 5 月にかけて実施され、スクリーニング後に 90 件の文献が採用された。定義、限界、受容性およびガイダンスに関するデータを抽出し、17 項目の試験報告を生成した。これら 17 項目を含む探索的レビューの結果は別報で公表済みである¹⁶⁾。プロジェクトチームでの議論を経て、13 項目を e-Delphi 調査での評価対象として選定した。

フェーズ 2 では、Core Outcome Measures in Effectiveness Trials (COMET) イニシアチブ (<https://www.comet-initiative.org/delphimanager/>) が管理する DelphiManager ソフトウェア (バージョン 5.0) を用

いて、2 ラウンドの e-Delphi 調査を実施し、潜在的な報告項目を 9 点リッカート尺度 (1-3: 重要でない, 4-6: 重要だが決定的ではない, 7-9: 決定的) で評価した。

第 1 回調査は 2022 年 8 月 24 日から 10 月 10 日まで、第 2 回調査は同年 10 月 31 日から 12 月 11 日まで実施された。参加者は以下の方法で特定された: 文献レビューから関連論文の著者への連絡、プロジェクトチームの専門的ネットワーク、学会および会議での参加者募集、ソーシャルメディア、専門組織およびネットワークを通じた周知 (Web appendix 2 に記載)。

合計 212 名の適格参加者が登録され、第 1 ラウンドでは 195 名 (92%)、第 2 ラウンドでは 176 名 (83%) が項目評価を実施した。参加者は 30 か国を代表し、試験研究者、試験方法論者 (統計学者を含む)、試験管理者、臨床医および関連医療専門家、代替指標の専門家、ジャーナル編集者、患者・一般市民パートナー、規制当局、支払者または医療技術評価専門家、倫理委員会および資金提供パネルメンバーなど、多分野にわたるステークホルダーで構成されていた。付録表 A2, A3, A4, A5 には参加者の特徴が記載されている (Web appendix 1)。

項目包含の閾値は 7-9 点スコア $\geq 70\%$ かつ 1-3 点スコア $< 15\%$ 、除外閾値は 1-3 点スコア $\geq 70\%$ かつ 7-9 点スコア $< 15\%$ とし、いずれの閾値も達成されない場合は包含・除外対象外とした。CONSORT-Surrogate の 13 項目が第 1 ラウンドで、14 項目が第 2 ラウンドで評価された (第 1 ラウンドで参加者から追加項目が提案された)。第 1 ラウンドで 8 項目、第 2 ラウンドでさらに 2 項目が閾値を達成した一方、両ラウンド終了後も閾値を達成しなかった項目は 4 項目であった (付録表 A6, Web appendix 1)。

第 3 フェーズでは、2023 年 3 月 13 日および 14 日に英国グラスゴーのグラスゴー大学および Zoom にて、ハイブリッド形式で合意形成会議が開催された。会議参加者には、プロジェクトチームメンバー 13 名と、e-Delphi 調査に参加したステークホルダー 20 名からなる招待サブグループが含まれた。e-Delphi 調査で閾値を達成しなかった 4 項目について議論と投票 (<https://www.mentimeter.com/>) が行われた。合意は「項目の包含 / 除外に 70% 以上が賛成」と事前定義された。4 項目全てで合意が形成

され、2項目が包含、2項目が除外された（付録表A7, Web appendix 1）。合意に達した項目については、会議出席者が文言の微調整や項目の統合を行い、e-Delphi 調査から提供された自由記述コメントについても議論された。

フェーズ4の主な目的は拡張版の項目の普及であり、現在も進行中である。普及活動には、プロジェクトを公表するための短編論文の出版³²⁻³⁶、プロトコルの出版^{30,31}、会議や学会での発表が含まれる。完成したチェックリストは、少なくとも1件の試験を実施した経験を持つ8名の試験研究者によりパイロットテストされた。具体的には、公表済みの試験を提供し、拡張項目が報告されているかどうかを注記するよう依頼した。全項目は明確であり、パイロットテストの結果に基づく変更は行われなかった。

CONSORT-Surrogate の構造

[Structure of the CONSORT-Surrogate extension]

CONSORT-Surrogate は、変更された項目または新規に追加された項目についての説明と、その根拠および解説を示すセクションを伴うチェックリストで構成されている。さらに、報告の具体例も提示されている。CONSORT-Surrogate において拡張されていない項目については、原版である CONSORT 2010 を参照する必要がある⁴⁾。

本ガイドラインでは、12件の公開試験報告書を用いて、11の CONSORT-Surrogate 項目それぞれについて少なくとも1つの報告例を提示した。例として用いられた試験報告のうち8件(67%)は、2017年1月から2022年6月までに6つの一般医学雑誌に掲載された試験プロトコルを対象としたターゲットレビューから特定されたものであり、残りは PubMed データベースを用いた探索的検索によって特定されたものである。

本稿の例文には、引用文献を示すために上付き文字の「ref」が付されている。また、いくつかの例には、CONSORT-Surrogate の利用を容易にする目的で、用語や推奨事項を補足している。必要に応じて、略語も例文中に記載している。

一方で、これらの報告例の提示は、当該試験の結果、結論、あるいは評価された介入を支持することを意味するものではない。さらに、CONSORT-Surrogate を適用すべき全ての疾患領域や研究分野

における事例を網羅的に特定し、列挙することは不可能である。しかしながら、研究チームはここでされた例を、それぞれの疾患領域や研究分野においてどのように報告すべきかを検討する際の指針として活用することができる。これらの例を理解することにより、試験報告書における CONSORT-Surrogate の適切な活用が促進されると考えられる。

対象を絞ったレビューにおける試験報告書の調査や事例収集など、多大な努力を払ったにもかかわらず、試験において代替エンドポイントが用いられていることを参加者に通知する取り組みを効果的に実施した例は確認されなかった。そこで、本 CONSORT-Surrogate の共著者である患者および公的機関のパートナー (DS, RH, SM, AW) と協力し、公開されている試験プロトコルからの引用を修正したうえで、この項目を試験報告書においてどのように活用できるかを示した (項目 26a.1)。

CONSORT-Surrogate

[CONSORT-Surrogate extension]

表3では CONSORT 2010 チェックリストと CONSORT-Surrogate チェックリストの拡張項目を比較した。付録3には、CONSORT 2010 と CONSORT-Surrogate のチェックリストをまとめたものが掲載されており、別途ダウンロードも可能である (Web appendix 3)。

タイトルと抄録 [Title and abstract]

[Items 1b (extended)]

CONSORT 2010 項目 1b [CONSORT 2010 item 1b]

試験デザイン、方法、結果、結論の構造化された要約⁴⁾

具体的なガイダンスについては、抄録の CONSORT を参照のこと³⁷⁾

CONSORT-Surrogate 項目 1b.1

[CONSORT-Surrogate extension item 1b.1]

1b.1 (a) 主要アウトカムが代替エンドポイントであること、および (b) 介入効果が代替される対象アウトカムを明記

表 3-1 CONSORT2010 と CONSORT-Surrogate の比較表

セクション	項目 No	CONSORT checklist	CONSORT-Surrogate
タイトルと抄録			
	1a	タイトルにランダム化試験であることを明記	—
	1b	試験デザイン、方法、結果、結論の構造化された要約（具体的なガイダンスについては、抄録の CONSORT を参照）	1b.1 (a) 主要アウトカムが代替エンドポイントであること、および (b) 介入効果が代替される対象アウトカムを明記
導入			
背景と根拠	2a	科学的背景と根拠の説明	2.1 (a) 主要アウトカムが代替エンドポイントであること、および (b) 介入効果が代替される目標アウトカムを明記
	2b	具体的な目的または仮説	
方法			
試験デザイン	3a	試験デザイン（並行群間比較、要因比較など）の説明（割付比を含む）	—
	3b	試験開始後の方法の重要な変更（適格基準など）とその理由	—
参加者	4a	参加者の適格基準	—
	4b	データ収集環境と場所	—
介入	5	各群に対する介入（実際の実施方法と時期、再現性を確保するための十分な詳細を含む）	—
アウトカム	6a	事前に規定された主要アウトカムと副次アウトカムの完全な定義（評価方法と時期を含む）	6a.1 代替エンドポイントを主要アウトカムとして使用する実際のまたは科学的な理由を記載 6a.2 選択された代替エンドポイントの正当性：(a) 代替エンドポイントの妥当性を示すエビデンス（あるいはその欠如）、および (b) 妥当性が使用された設定および状況（例：介入、疾患、集団）に特有のものであることを示すエビデンス（あるいはその欠如）
	6b	試験開始後のアウトカムの変更（理由を含む）	—
サンプルサイズ	7a	サンプルサイズの決定方法	7a.1 代替エンドポイントの最小の効果量から目標アウトカムでの有効性が予測されることを示すために、サンプルサイズが推定されたかどうかを明記
	7b	該当する場合、中間解析および中止ガイドラインの説明	—
ランダム化：			
シーケンス生成	8a	ランダム割付シーケンスを生成するために使用された方法	—
	8b	ランダム化の種類、制限事項（ブロッキングやブロックサイズなど）の詳細	—
ランダム割付シーケンス	9	ランダム割付シーケンスを実行するために使用されたメカニズム（連番の付いたコンテナなど）、介入が割り当てられるまでシーケンスを秘匿するために行われた手順の説明	—
実施	10	ランダム割付シーケンスを生成した者、参加者を登録した者、参加者を介入に割り当てた者	—
盲検化	11a	介入への割り当て後に盲検化された者（例：参加者、ケア提供者、アウトカム評価者）とその方法	—
	11b	関連する場合、介入の類似性の説明	—
統計学的手法	12a	主要アウトカムと副次アウトカムのグループ比較に使用された統計学的手法	—
	12b	サブグループ解析や調整解析などの追加解析の方法	—

表 3-2 CONSORT2010 と CONSORT-Surrogate の比較表

セクション	項目 No	CONSORT checklist	CONSORT-Surrogate
結果			
参加者フロー (図示を強く推奨)	13a	各群について、ランダムに割り付けられ、意図された治療を受け、主要アウトカムについて解析された参加者数	—
	13b	各群について、ランダム化後の脱落および除外例とその理由	—
募集	14a	募集期間および追跡期間を定める日付	—
	14b	試験が終了または中止された理由	—
ベースライン データ	15	各群のベースラインの人口統計学的および臨床的特徴を示す表	—
解析対象者	16	各群について、各解析に含まれた参加者数(分母)と、解析が当初の割り付け群によるものかどうか	—
結果と推定	17a	各主要アウトカムおよび副次アウトカムについて、各群の結果、推定効果量とその精度(95%信頼区間など)	17a.1 主要アウトカムが代替エンドポイントを含む複合アウトカムである場合は、複合アウトカムの全てのコンポーネントに対する介入効果を報告
	17b	二値アウトカムについては、絶対効果量と相対効果量の両方を提示することが推奨される	—
補助解析	18	サブグループ解析や調整解析など、実施されたその他の解析の結果(事前に規定された解析と探索的解析を区別する)	—
害	19	各群における全ての重要な有害事象または意図しない影響(具体的なガイダンスについては、有害事象に関する CONSORT を参照)	—
議論			
研究限界	20	試験の限界、潜在的なバイアス、不正確さ、そして関連する場合は分析の多重性の原因への対処	—
一般化	21	試験結果の一般化可能性(外的妥当性、適用可能性)	—
解釈	22	解釈は結果と一致しており、ベネフィットと害のバランスを取り、他の関連するエビデンスを考慮しているか	22.1 代替主要アウトカムを用いた場合の試験結果の解釈(目標アウトカムに対する介入効果の既知の妥当性、および参加者に対する試験介入の潜在的なベネフィット・リスク評価を含む)
			22.2 代替エンドポイントの使用を考慮し、試験デザイン(サンプルサイズおよび追跡期間を含む)が、試験対象者の介入の潜在的な危害を適切に捉えているかどうかの評価
			22.3 目標アウトカムに関する現在の知見を検証するために、今後実施する分析/研究の計画についての記述
その他の情報			
試験登録	23	試験登録番号および名称	—
プロトコル	24	試験プロトコル全文にアクセスできる場所(入手可能な場合)	—
資金	25	資金源およびその他の支援(薬剤の供給など)、資金提供者の役割	—
	26	新項目	26.1 試験参加者が登録前に、試験が代替エンドポイントを用いて介入の効果を評価するように設計されていることを知らされたかどうか、またどのように知らされたかを明記 26.2 試験において代替エンドポイントおよび目標アウトカムのデータが収集された場合、将来の二次研究のためにデータのオープンアクセスの取り決めを明記

付録 3 には、CONSORT 2010 と CONSORT-Surrogate のチェックリストをまとめたものが掲載されており、別途ダウンロードも可能である(Web appendix 3)。 CONSORT = Consolidated Standards of Reporting Trials.

CONSORT-Surrogate 項目 1b.1 の例

[Examples of CONSORT-Surrogate item 1b.1]

例 1 [Example 1]

「主要アウトカムは、腎障害の代替マーカー [エンドポイント] である尿中好中球ゼラチナーゼ関連リポカリンの 48 時間以内のピーク変化であった³⁸⁾。」(「エンドポイント」という語句を追加、その使用を推奨する)

例 2 [Example 2]

「早期 2 型糖尿病患者における動脈硬化および早期アテローム性動脈硬化の代替 [エンドポイント] マーカーとして、ジペプチジルペプチダーゼ-4 (DPP-4) 阻害剤リナグリプチンが大動脈脈波伝播速度 (PWV) に及ぼす影響を評価する³⁹⁾。」(「マーカー」ではなく「代替エンドポイント」という用語の使用を推奨する)

解説 [Explanation]

適切に作成された試験抄録は、読者が報告書全文を読むか、あるいはアクセスするかを判断するための初期的な評価を提供するものであり、場合によっては医療上の意思決定の根拠となることもある^{37, 40)}。特定の項目に明示的に言及した抄録は、データベースにおける索引付けや³⁷⁾、それに伴う代替エンドポイントの検証などの二次研究のために試験を検索する際に重要である。

しかしながら、その重要性にもかかわらず、スペースの制約により、抄録には試験の主要な情報のみを記載する必要がある⁴¹⁾。したがって、抄録の作成にあたっては CONSORT の指針を用いることに加え³⁷⁾、試験著者は、代替エンドポイントが主要アウトカムとして使用されていること、およびそれが代替する目標アウトカムについて明示的に記載する必要がある。なお、抄録の構成は多様であり、利用可能なスペースも限られている^{40, 41)}。そのため、具体的な記載方法については、提示されている例を参照されたい。

導入 [Introduction]

背景と根拠 [Background and objectives (extended)]

CONSORT 2010 項目 2a [CONSORT 2010 item 2a]

科学的背景と根拠の説明⁴⁾

CONSORT を参照³⁷⁾

CONSORT 2010 項目 2b [CONSORT 2010 item 2b]

具体的な目的または仮説⁴⁾

CONSORT を参照³⁷⁾

CONSORT-Surrogate 項目 2.1

[CONSORT-Surrogate extension item 2.1]

1b.1 (a) 主要アウトカムが代替エンドポイントであること、および (b) 介入効果が代替される対象アウトカムを明記

CONSORT-Surrogate 項目 2.1 の例

[Example of CONSORT-Surrogate extension item 2.1]

「PWV (大動脈脈波伝播速度) は動脈の機能と構造の総合的な指標であり、したがって早期アテローム性動脈硬化の [代替エンドポイント] マーカーである^{ref)}。PWV が高いほど、動脈硬化が進行し、心血管イベント (CV) のリスクが高まる^{ref, 39)}。」(例には「代替」という語を追加しており、該当項目を報告する際にはこの語の使用を推奨する。また、「代替エンドポイント」という用語の使用と、代替エンドポイントの妥当性を裏付ける参考文献の具体的な引用も推奨される (項目 6a.2 の解説を参照))

解説 [Explanation]

序論では、最新のエビデンスおよび解決すべき知識ギャップを要約することにより、試験実施の理由を概説する必要がある^{4, 40)}。序論の内容および構成に関する詳細については、REPORT ガイドを参照されたい⁴⁰⁾。導入部は読者に試験報告書の概要を提示する役割を果たし^{4, 40)}、同時にジャーナル編集者や査読者が当該試験報告の重要性を評価する機会ともなる⁴⁰⁾。

したがって、試験著者は、代替エンドポイントの使用および治療効果の代替指標となる目標アウトカムについて明示的に記述する必要がある。最終的な試験報告論文の導入セクションは、試験プロトコルの論文よりも短くなる場合があるため⁴¹⁾、本項目では、主要アウトカムが代替エンドポイントであることと、それに対応する目標アウトカムについて簡潔に記述するだけで十分である。なお、試験著者は方法セクションにおいて、選択した代替エンドポイ

ントの詳細およびその根拠を説明することができる(項目 6a.1 および 6a.2 を参照)。

一方で、読者が試験の背景や重要性をより適切に理解できるよう、これらの内容を導入部で要約して示すことも可能である。最後に、導入部の構成や語数はジャーナルごとに異なるため、試験著者は CONSORT 2010 の項目 2a または項目 2b のいずれかを報告する際に、本項目の内容を併せて記載することができる。

方法 [Methods]

アウトカム [Outcomes]

CONSORT 2010 項目 6a

[CONSORT 2010 item 6a (extended)]

事前に規定された主要アウトカムと副次アウトカムの完全な定義(評価方法と時期を含む) CONSORT⁴⁾ と CONSORT-Outcomes⁷⁾ を参照

CONSORT-Surrogate 項目 6a.1

[CONSORT-Surrogate extension item 6a.1]

代替エンドポイントを主要アウトカムとして使用する実際のまたは科学的な理由を記載

CONSORT-Surrogate 項目 6a.2

[CONSORT-Surrogate extension item 6a.2]

選択された代替エンドポイントの正当性: (a) 代替エンドポイントの妥当性を示すエビデンス(あるいはその欠如)、および (b) 妥当性が使用された設定および状況(例: 介入, 疾患, 集団)に特有のものであることを示すエビデンス(あるいはその欠如)

CONSORT-Surrogate 項目 6a.1 の例

[Example of CONSORT-Surrogate item 6a.1]

「本試験では、試験費用、この分野における急速に進展するエビデンス、そして脆弱な集団における長期介入の実施可能性に関する懸念など、いくつかの実践的な制約のため、代替エンドポイントを使用した。しかしながら、選択されたエンドポイントは、心血管(CVD) イベントの予後予測に有意であることが検証されている⁴²⁾。」

解説 [Explanation]

代替エンドポイントに関連する制約¹⁶⁾を踏まえ、

試験著者は代替エンドポイントを使用する科学的または実務的な理由を読者に説明する必要がある。代替エンドポイントを用いる理由として最も一般的に挙げられるのは、試験の効率性の向上、すなわち追跡期間の短縮や必要サンプルサイズの削減である。このような利用は、生物学的活性の実証や目標アウトカムに基づく将来の試験の必要性に関する情報提供に重点を置く初期段階の試験に特に適している¹⁶⁾。

また、一次予防試験では症例の集積に長い時間を要する場合があります、希少疾患を対象とする試験では利用可能な試験集団が小規模に限られることが多い¹⁶⁾。さらに、代替エンドポイントは、重篤で生命を脅かす疾患において未充足の医療ニーズが高い場合、迅速承認または加速承認の枠組みの一環として、規制当局の承認手続きにおいて広く用いられている^{8, 16, 43)}。加えて、介入の状況によっては目標アウトカムの測定が必ずしも適切ではない場合もある。例えば、小児試験では参加者報告アウトカムの使用が、新生児や非常に若い小児(7歳未満)には困難であり、その場合には観察者報告アウトカムが必要となる可能性がある⁴⁴⁾。なお、ここで示した代替エンドポイントを使用する実務的または科学的理由¹⁶⁾は、必ずしも網羅的なものではない。

本項目を報告することにより、読者は代替エンドポイントを主要アウトカムとして用いる正当性を理解することができ、試験の重要性を文脈に沿って把握することが可能となる。ただし、本項目を適切に報告した場合であっても、試験著者が項目 6a.2 において、選択した代替エンドポイントの妥当性について検討・議論することを妨げるものではない(項目 6a.2 の解説を参照)。

CONSORT-Surrogate 項目 6a.2 の例

[Example of CONSORT-Surrogate item 6a.2]

例 1 [Example 1]

「これらの試験の主要アウトカムは、米国食品医薬品局(FDA)との合意に基づき選択された。副甲状腺ホルモン(PTH)の30%以上の減少に関連する明白な臨床的ベネフィットを具体的に実証した公表データは見つからなかったが、いくつかの観察研究において、[代替評価項目として] PTH濃度が600 pg/mLを超える場合、PTH濃度が150~300

pg/mLの場合と比較して、[目標アウトカムである]死亡、心血管イベント、骨折の発生率が高くなることが示されている^{ref. 45)}。」(角括弧内の文言を追加した。該当項目の報告時には使用を推奨する)

例 2 [Example 2]

「主要有効性アウトカムは、ベースラインから2か月間の日中歩行時収縮期血圧の変化であった。収縮期血圧は、心血管イベントおよび死亡率の予測における有効な代替エンドポイントである。これは、123件の降圧薬試験のメタアナリシスに基づくもので、613,815人の参加者において、収縮期血圧の治療効果と心血管イベントの間に強い関連性が示された^{ref)}。具体的には、メタ回帰分析により、主要な心血管疾患イベント ($P < 0.0001$)、脳卒中 ($P < 0.0001$)、心不全 ($P < 0.0001$)、および全死亡率 ($P = 0.014$) の相対リスク低下は、達成された収縮期血圧低下の程度に比例することが示された。しかし、様々な疾患のリスク低下は薬剤クラス間で異なり、腎神経遮断術を用いた場合でも、血圧低下が心血管イベントおよび死亡率のベネフィットを予測する妥当性を確立するには、さらなるエビデンスが必要である。」(この例は、著者が既に公表されている試験¹⁰⁾を基に作成したものであり、当該試験で引用されているメタアナリシス⁴⁶⁾を用いて構成したものである。メタアナリシス⁴⁶⁾では、血圧降下薬による介入を評価したランダム化比較試験において、代替エンドポイント(収縮期血圧の差)と目標アウトカム(心血管疾患および全死亡の相対リスク)に対する治療効果との間に強い関連があることが報告されている。詳細については、項目 6a.2 の解説を参照されたい)

解説 [Explanation]

代替エンドポイントは、使用に先立って検証される必要がある。ここでいう検証とは、介入が代替エンドポイントに及ぼす影響が、目標アウトカムに対する介入効果をどの程度予測できるかを評価することである^{47, 48)}。代替エンドポイントの検証に関する詳細な議論は、本拡張ガイドラインの範囲を超えるものであるが、読者に対しては、代替エンドポイントの検証方法に関する複数の論文⁴⁷⁻⁵⁶⁾、エビデンスの妥当性を評価するためのフレームワーク^{21, 57-59)}、

さらに近年提案された代替エンドポイント検証の報告のためのチェックリスト⁶⁰⁾を参照することを推奨する。要点として、代替エンドポイントの検証では、代替エンドポイントと目標アウトカムとの間に強い関連が存在すること(いわゆる個人レベルの関連)と、代替エンドポイントに対する治療効果が目標アウトカムに対する治療効果と密接に関連していること(いわゆる試験レベルの関連)の双方が示される必要がある^{47, 48)}。

例えば、項目 6a.2 の例 1 は、副甲状腺ホルモン(PTH: 代替エンドポイント)と死亡率および臨床イベントという目標アウトカムとの観察研究に基づく関連性に依拠しており、望ましいエビデンスレベルには達していない。これに対して例 2 では、ランダム化比較試験のメタアナリシス(回帰分析)に基づき、収縮期血圧という代替エンドポイントと死亡率という目標アウトカムに対する治療効果の関連性が引用されている。

代替エンドポイントの妥当性検証の強さを適切に評価するためには、試験著者は以下の主要なメタ回帰指標を提示すべきである。すなわち、代替エンドポイントに対する治療効果と目標アウトカムに対する治療効果との線形関係における傾き係数(および95%信頼区間)、スピアマンの相関係数(ρ)や R^2 などの関連性の強さを示す指標、さらに代替治療効果(surrogate threshold effect)または予測区間である(項目 7a.1 を参照)。血圧と心血管イベントに関するこれらの指標の具体例については、Lassere らによる論文⁵⁸⁾に示されている。

試験における代替エンドポイントの妥当性検証は、より適切に報告される必要がある。2005年および2006年に公表された626件の試験を分析した研究では、主要アウトカムとして代替エンドポイントを使用していた試験報告のうち、その妥当性について議論していたものはわずか34%(109件中37件)に過ぎなかったことが示されている⁶¹⁾。

がん領域においては、代替エンドポイントの妥当性に関する複数の研究が報告されている。あるシステマティックレビューでは、代替エンドポイントの妥当性は全体として比較的低いことが示された。具体的には、試験で使用された代替エンドポイントの52%は、治療効果と全生存率という目標アウトカムとの相関が低く($r \leq 0.7$)、高い相関($r \geq 0.85$)

を示したものはわずか23%であった。

代替エンドポイントの検証モデルは、代替エンドポイントに対する効果が推定されている新規試験において、目標アウトカムに対する治療効果を予測する機会を提供することが多い。そのため、予測の精度を定量化することが重要である⁶⁰⁾。モデルの予測性能および較正を評価するためには、leave-one-outクロスバリデーションに加え、モデル構築後に公表された新たな試験や、モデル推定時に個々の患者データが利用できなかった試験を用いた外部検証が不可欠である⁶²⁾。

したがって、試験著者は、代替エンドポイントの妥当性に関するエビデンス（あるいはその欠如）を明確に提示する必要がある。これまでに、代替エンドポイントの妥当性検証のための多くの統計的アプローチが提案されてきており^{54,63)}、その一部はボックス2に要約されている。したがって、試験著者は、代替エンドポイントの選択の根拠となるアプローチについても詳細に提示する必要がある。

ボックス2

代替エンドポイントの妥当性検証のための統計的アプローチの概要

治療効果評価における代替エンドポイントの妥当性を評価するための、過去40年間に登場した、厳選された非網羅的な統計的手法と一般的なアプローチ

プレンティス基準⁵³⁾

1989年に発表された先駆的な研究において、プレンティスは仮説検定の外挿の妥当性に関する3つの基準を提案した（代替エンドポイントへの治療効果がないという帰無仮説を棄却することは、目標アウトカムへの治療効果がないという帰無仮説を棄却することを意味する）。

- 代替エンドポイントが真のエンドポイントに及ぼす影響は、ランダム化群によって変化しない。
- 代替エンドポイントが真のエンドポイントに影響を及ぼす。
- 治療が代替エンドポイントに及ぼす影響は、治療が真のエンドポイントに及ぼす平均効果を変化させる。

プレンティス基準は概念的には依然として重要であるが、実践においては有用性が限られている。

主層別化⁶⁴⁾

この手法では、因果効果が代替エンドポイント評価の基礎となるべきであると主張している。因果効果とは、同一の個人集団における潜在的なアウトカムの治療群間の比較である。代替妥当性には2つの要件が必要である。1つは因果的必然性であり、治療が目標アウトカムに与える影響は、治療が代替アウトカムにも影響を与えた場合にのみ存在することを要求する。もう1つは統計的一般化可能性であり、代替アウトカムのみが観察される将来の研究において、代替アウトカムが目標アウトカムに対して良好な予測性能を示すことを要求する。

メタ分析的回帰に基づくアプローチ^{47,65)}

このアプローチは、多重試験（ランダム化試験）設定における代替アウトカムと目標アウトカムの2段階の共同モデリングに基づく。代替効果は、個々の患者レベルにおける代替アウトカムと目標アウトカム間の決定係数（個人レベル R^2 ）、および試験レベルにおける代替アウトカムと目標アウトカム間の決定係数（試験レベル R^2 ）に基づいて確立される。あるいは、代替アウトカム閾値効果は、目標アウトカムにも有意な治療効果が存在すると結論付けるために必要な、代替アウトカムに対する最小の治療効果レベルを定義する実用的な尺度として提案されている⁶⁶⁾。情報理論に基づくこれらのメタ分析手法の拡張は、因果関係パラダイムにおける推奨アプローチとして提案されている⁶⁷⁾。

ベイズ的アプローチ

ベイズ的アプローチは、上述した全ての方法論に比較的容易に適用することが可能である。なかでも最も一般的に用いられるモデルとしては、DanielsとHughes⁶³⁾によって提案されたメタアナリシスに基づく固定（独立）効果モデルと、目標アウトカムおよび代替エンドポイントに対する試験レベルの効果を同時にモデル化するベイズ的ランダム効果メタアナリシスが挙げられる⁵¹⁾。さらに近年では、代替エンドポイントと目標アウトカムに対する治療効果の関連性を明示的に考慮するベイズ多変量メタアナリシスが提案されており、とくに規制当局による承認判断や医療技術評価における償還(reimbursement)の意思決定の文脈での活用が期待されている⁵¹⁾。

さらに、ある試験環境（例：十分に類似した対象集団，介入，疾患，対照群，設定）における代替指標の妥当性のエビデンスは，別の環境には一般化できない可能性がある¹⁶⁾。例えば，無増悪生存期間を全生存期間の代替エンドポイントとして評価した研究のシステマティックレビューでは，試験レベルの妥当性が評価対象の介入，がんの局在，病期によって異なっていたことが示されている⁶⁸⁾。

罹患率や死亡率の改善を予測する体格指数（BMI）を用いて評価される体重減少の程度は，疾患や肥満に関連する合併症，個人の年齢，およびベースラインの肥満レベルによって異なることが多い^{69, 70)}。したがって，試験責任医師は，試験の環境における代替有効性のエビデンス（あるいはその欠如）に基づいて代替エンドポイントを正当化すべきである。有効性に関する項目 6a.2 の例 2 では，異なる疾患に特化しているものの，試験対象の特定の介入に関するエビデンスの欠如を認めていることが示されている。

サンプルサイズ [Sample size]

CONSORT 2010 項目 7a

[CONSORT 2010 item 7a (extended)]

サンプルサイズの決定方法
CONSORT4 を参照

CONSORT-Surrogate 項目 7a.1

[CONSORT-Surrogate extension item 7a.1]

代替エンドポイントの最小の効果量から目標アウトカムでの有効性が予測されることを示すために，サンプルサイズが推定されたかどうかを明記

CONSORT-Surrogate 項目 7a.1 の例

[Example of CONSORT-Surrogate item 7a.1]

例 1 [Example 1]

「以前発表されたデータから，当センターにおける造影剤関連急性腎障害（CA-AKI）の全体的な発生率が低いことが示唆されたため，主要アウトカムとして中性球ゼラチナーゼ関連リポカリン（NGAL）を選択した。本探索的研究の主要アウトカムについては，研究設計時点で rhCIINH を用いた予防療法研究に関する適切なデータが不足していたため，検出力の正式な計算は行われなかった。したがって，

検出力計算には不十分な推定値が使用される可能性があった。過去の異なる予防レジメン^{ref.5)}を用いた介入研究および類似の腎機能代替指標を参考に，標準偏差 150 ng/mL，検出力 80%，両側第 1 種誤り 5% を仮定し，平均尿中ピーク NGAL 濃度 100 ng/mL の差を検出可能とするには各群 40 例が必要と算出した。この差は [目標アウトカムである] 急性腎障害（AKI）の予測因子であることが示されている^{ref. 38)}。」（角括弧内の語句を追加。本試験の探索的性質を考慮し，著者は観察研究の指標を使用しているが，試験チームは代替指標の妥当性検証研究から直接得られた指標の使用を推奨する）

例 2 [Example 2]

「検出力計算の前提条件（6 分間歩行試験距離における最小臨床的意義のある改善として 40 m 増加を閾値とする [代替エンドポイント閾値効果]，SD [標準偏差] は 80 m とする）は，以下の 2 点に基づいた：(1) 肺動脈性高血圧症患者を対象とした過去のランダム化臨床試験のメタ回帰分析（HFpEF [駆出率保持型心不全] 患者における同データ不足のため）^{ref.6)} および (2) 試験運営委員会のメンバーによる臨床的合意⁷¹⁾。」（引用された代替指標検証研究と整合性を保つため，「最小臨床的意義のある改善」ではなく「代替閾値効果」という用語の使用を推奨する）

解説 [Explanation]

試験のサンプルサイズ決定は，目標効果サイズの詳細や結果における試験脱落の許容度を含め，適切に正当化され，十分に報告されなければならない^{4, 72)}。主要アウトカムが代替エンドポイントである試験では，代替有効性の指標に基づいて目標効果サイズの選択を検討すべきである。例えば，よく報告される妥当性指標として，目標アウトカムにおける治療効果を予測するために必要な，代替エンドポイントに対する最小治療効果（代替閾値効果として知られる）がある^{58, 66)}。代替閾値効果の概念は項目 7a.1 の例 2 で使用されたが，著者らはこれが異なる患者集団から導出されたものであることを認めている（彼らの試験集団のデータが利用不可能であったため）⁷¹⁾。これに対し，項目 7a.1 の例 1 では³⁸⁾，他の代替有効性指標が用いられている。

すなわち目標アウトカムを予測する代替エンドポイントの差異の正当化は、受信者動作特性曲線（ROC曲線）から導出されたカットオフ閾値を用いた前向き研究に基づいている⁷³⁾。

一部の事例では、事前の代替エンドポイントの妥当性検証が行われていないため、著者が代替有効性指標を用いて正式に検出力（サンプルサイズ）を決定できない場合がある。ただし、データが入手可能な場合、試験研究者は選択した代替エンドポイントの事前の妥当性検証を検討できる（項目 26.2 参照）。さらに、代替エンドポイントは主に試験効率の向上（すなわち、目標アウトカムを使用する場合と比較してより少ないサンプルサイズを可能にする）のために使用されることを考慮すると、試験著者は代替エンドポイントと目標アウトカムの両方についてサンプルサイズを決定することが推奨される。目標アウトカムに対する治療効果に基づくサンプルサイズが、代替エンドポイントに基づくサンプルサイズと同等（またはそれ以下）である場合、代替エンドポイントを主要アウトカムとして選択した十分な根拠を示すべきである。

最後に、妥当性指標の使用の有無にかかわらず、試験著者は代替エンドポイントの使用とその既知の

妥当性（項目 22.1 参照）の文脈における結果の解釈について議論すべきである。これには、目標アウトカムに対する予測効果とその不確実性（信頼区間によって反映される）がどのように導出されたかも含まれる。

結果 [Results]

結果と推定 [Outcomes and estimation]

CONSORT 2010 項目 17a [CONSORT 2010 item 17a (extended)]

各主要アウトカムおよび副次アウトカムについて、各群の結果、推定効果量とその精度（95% 信頼区間など）

CONSORT⁴を参照

CONSORT-Surrogate 項目 17a.1

[CONSORT-Surrogate extension item 17a.1]

主要アウトカムが代替エンドポイントを含む複合アウトカムである場合は、複合アウトカムの全てのコンポーネントに対する介入効果を報告

CONSORT-Surrogate 項目 17a.1 の例

表 4 および表 5 を参照

表 4 複合アウトカムのコンポーネントの報告：米国リウマチ学会 (ACR) レスポンス。表は van de Putte ら⁷⁴⁾の結果を用いて作成し、BMJ Publishing Group の許可を得て掲載

測定項目	アダリムマブ 40 mg / 2 weeks (n=113)	プラセボ (n=110)
ACR レスポンス (%)		
ACR 20	46.0	19.1
ACR 50	22.1	8.2
ACR 70	12.4	1.8
個別指標—絶対変化率 (%)		
圧痛関節数 (0-68)	-13.6 (-37.4)	-6.6 (-9.5)
腫脹関節数 (0-66)	-8.5 (-37.0)	-2.4 (-7.4)
患者による疼痛評価 (VAS 0-100)	-27.6 (-37.7)	-11.0 (-11.4)
患者による疾患活動性の全般評価 (VAS 0-100)	-27.9 (-38.9)	-10.6 (-7.9)
医師による疾患活動性の全般的評価 (VAS 0-100)	-27.3 (-38.8)	-10.9 (-12.9)
健康評価質問票 (機能評価) の変化	-0.38 (-21.3)	-0.07 (+1.8)
C 反応性タンパク質 (mg/mL)	-19.5 (-42.8)	+3.0 (+0.4)
赤血球沈降速度 (mm/h)	-12 (-28.8)	-2.0 (-4.4)

特に断りがない限り、データは平均値（標準偏差）で示される。

ACR20 は複合指標であり、圧痛関節数および腫脹関節数の 20% 改善、ならびに以下の 5 基準のうち 3 項目における 20% 改善を定義する：患者全体評価、医師全体評価、機能能力測定（通常は健康評価質問票）、疼痛の視覚的アナログ尺度 (VAS)、および代替エンドポイント（赤血球沈降速度または C 反応性タンパク質）。ACR50 および ACR70 は ACR20 と同様の指標であるが、改善レベルがそれぞれ 50% および 70% と定義される。

表5 Per protocol population (PPS) における複合アウトカム（無増悪生存期間）の各コンポーネント（再発または死亡を含む）の報告。表は Parekh ら⁷⁵⁾ を改変し、エルゼビアおよび著作権クリアランスセンターの許可を得て転載

項目	ロボット支援膀胱全摘除術 (n=150)	開腹膀胱全摘除術 (n=152)
総イベント数	49 (32.7)	50 (32.9)
膀胱がんによる死亡	28 (18.7)	32 (21.3)
非がん性死亡	10 (6.7)	11 (7.2)
再発（最終連絡時生存中）	11 (7.3)	7 (4.6)
あらゆる再発	39 (26.0)	39 (26.3)
純粋な局所再発（総数）*	6 (4.0)	4 (2.6)
膀胱摘出床*	6 (4.0)	2 (1.3)
骨盤リンパ節郭清部位	0	1
腹壁またはポートサイト	0	1
遠隔再発（局所再発の有無を問わず；合計）	33 (22.0)	35 (23.0)
肺	8	10
肝臓	6	7
骨	9	10
骨盤外リンパ節	9	9
腹膜播種	2	1
副腎	2	1
結腸	—	3
小腸	1	—
腎臓	1	—
脳	—	1
不明	4	4
二次性尿路上皮がん（合計）	1 (0.6)	3 (2.0)
上部尿路	1	2
尿道	0	1
検閲観察（総数）	101 (67.3)	102 (67.1)
2年以内に検閲	10 (6.7)	14 (9.2)
経過観察中	9	8
経過観察不能	1	6

データは人数 (%) で示す。疾患の進行は、固形腫瘍の治療効果評価基準 (RECIST) バージョン 1.1 に基づき、画像所見または病理学的所見による疾患のエビデンス（代替エンドポイント）、あるいは疾患による死亡（主要アウトカム）によって定義された。文書化された再発または他の原因による死亡も進行とみなした。

* 全局所再発率 ($P=0.54$) および膀胱切除床再発率 ($P=0.17$) において、群間差は認められなかった。

解説 [Explanation]

複合アウトカムは、2つ以上のコンポーネントアウトカム（例：収縮期血圧上昇、非致死性脳卒中発症、死亡のいずれかを経験した参加者の割合）で構成される。いずれかのコンポーネントアウトカムにカウントされた場合は、複合アウトカムにもカウントする²⁵⁾。試験における複合アウトカムの使用に関する考慮事項は、本拡張の範囲外であり、別稿で詳細に論じられている^{25, 76-78)}。

2008-10年に発表された試験の監査では、複合アウトカムを用いた106試験のうち28% ($n=30$) がコンポーネントの一つとして代替エンドポイントを

含んでいたことが判明した⁷⁹⁾。試験著者らは、複合アウトカムの各コンポーネントに対する治療効果を個別に報告することが推奨される。本項目の報告は、明示的な複合アウトカム（例：「主要アウトカムは複合アウトカムであった [コンポーネント a または b または c]」）および複合測定（表4および表5）に適用される。複合指標および尺度には、がんにおける疾患無増悪生存期間（腫瘍サイズによる疾患再発・進行の測定、または死亡）⁸⁰⁾、感染症における臨床治癒（臨床医評価による反応、画像所見または微生物学的基準による測定）^{81, 82)} など、代替エンドポイントを含む複合アウトカムが組み合わされ

る。本項目は、補足ファイルや付録ではなく、試験報告の本文中に記載することが望ましい。

議論 [Discussion]

解釈 [Interpretation]

CONSORT 2010 項目 22 (拡張)

[CONSORT 2010 item 22 (extended)]

解釈は結果と一致しており、有益性と有害性のバランスを取り、他の関連するエビデンスを考慮している。

CONSORT 20104 を参照。追加された3つのCONSORT-Surrogate に準じた報告は、まとめて、かつ任意の順序で行うことができる。

CONSORT-Surrogate 項目 22.1

[CONSORT-Surrogate extension item 22.1]

代替主要アウトカムを用いた場合の試験結果の解釈（目標アウトカムに対する介入効果の既知の妥当性、および参加者に対する試験介入の潜在的な有益性・リスク評価を含む）

CONSORT-Surrogate 項目 22.2

[CONSORT-Surrogate extension item 22.2]

代替エンドポイントの使用を考慮し、試験デザイン（サンプルサイズおよび追跡期間を含む）が、試験対象者の介入の潜在的な害を適切に捉えているかどうかの評価

CONSORT-Surrogate 項目 22.3

[CONSORT-Surrogate extension item 22.3]

目標アウトカムに関する現在の知見を検証するために、今後実施する分析/研究の計画についての記述

CONSORT-Surrogate 項目 22.1 の例

[Example of CONSORT-Surrogate item 22.1]

例 1 [Example 1]

「浸潤性乳がんの検出率は早期スクリーニングの代替指標であり、TOSYMA の結果 [試験名] はデジタル乳房トモシンセシスが長期的なスクリーニング効果に及ぼす可能性のある影響を示唆している。TOSYMA のスクリーニング段階における早期腫瘍段階の浸潤性乳がん検出率の絶対的増加は、おそらく診断の改善を示しており、スクリーニング対象集

団における進行性乳がんの発生率を低下させ、ひいては乳がん死亡率にプラスの効果をもたらす可能性が期待される^{ref5)}。しかし、スクリーニングで小さながんがより多く発見されても、次のスクリーニング検査までの2年間にスクリーニング陰性女性の間で浸潤性間期がんの発生率が減少しない場合、スクリーニングの利点と過剰診断の可能性について疑問が生じるだろう^{refs, 83)}。」

例 2: 22.1 と 22.2 の統合項目

[Example 2: Combining items 22.1 and 22.2]

「心血管イベント高リスクの難治性高血圧患者における腎神経切断術後、我々が観察した診察室収縮期血圧 [代替エンドポイント] の平均 9.0 mmHg 低下は^{ref)}、Global SYMPPLICITY レジストリ^{ref)} の3年報告および RADIANCE-HTN SOLO 試験^{ref)} の12か月結果が示すように長期に維持される場合、降圧薬治療における [目標アウトカム:] 脳卒中、冠動脈疾患、心不全、全死因死亡率の減少と関連付けられてきた効果と同等の大きさである^{ref)}。腎神経切断術後の血圧変動幅が減少するという RADIANCE-HTN SOLO 試験における我々の先行観察結果^{ref, 10)}を確認できれば、心血管イベントと脳血管イベントの両方の減少も期待できる。」（角括弧内の文言を追加した。該当項目の報告時には使用を推奨する）

解説 [Explanation]

本項目では、試験著者が使用した特定の代替エンドポイントの主な限界を考慮し、試験結果の解釈への影響について議論することを推奨する。具体的には、読者は以下の点について知らされるべきである：代替エンドポイントで観察された介入効果に対する結果の意味、現在の妥当性エビデンス（あるいはその欠如）に基づく目標アウトカムへの影響、および潜在的な介入の全体的な有益性と有害性のバランス。関心領域において代替エンドポイント検証モデルの予測性能が良好な場合、研究者は目標アウトカムに対する予測効果を、不確実性の尺度（例：信頼区間）および実際に使用した予測式とともに提示すべきである。以下の代替エンドポイント検証研究はこのレベルの報告例を示している⁸⁴⁻⁸⁶⁾。本項目の報告は、関係者が試験結果を別の研究の実施や方針の指針として活用する方法を伝える上で重要である。

より長期の追跡調査と大規模なサンプルサイズで代替エンドポイントと目標アウトカムの両方のデータを収集する試験は、介入の有益性と有害性のバランスについて推測の余地が少ない。標本サイズや追跡期間にかかわらず、試験実施者は収集した対象アウトカムに対する治療効果を報告すべきである。代替有効性に関するその他の拡張項目（項目 6a.2）および潜在的な有害事象（項目 22.1）について適切に報告することで、本項目の適切な報告が可能となる。

CONSORT-Surrogate 項目 22.2 の例

[Example of CONSORT-Surrogate item 22.2]

「追加の追跡調査が必要であり、超音波腎神経切断術による降圧効果が時間経過とともに持続するかどうか、特に患者が追加の降圧薬（特にアルドステロン拮抗薬スピロラクソン）を投与された場合、盲検下（2~6 か月）および非盲検下（6 か月以降）の両条件下で血圧をコントロールできるかどうかを判断する必要がある^{ref)}。有害事象は稀であったものの、本試験の追跡期間延長と治療対象患者の増加により、追加の安全性データを提供する必要があり¹⁰⁾。」

解説 [Explanation]

代替エンドポイントに対する試験的治療効果は介入の潜在的な有益性を示唆する可能性がある一方で、長期的な試験追跡調査や日常診療への導入により、その介入が有害であることが明らかになる場合もある⁸⁷⁾。1996年、フレミングとデメッツは、代替エンドポイントに対する治療効果の陽性結果に基づいて承認された薬剤が、後に患者や公衆全体に害を及ぼすことが判明した事例を複数報告した⁸⁸⁾。この例には不整脈（異常な心拍リズム）の抑制が含まれ、心血管関連死亡の代替エンドポイントと見なされた不整脈を減少させる薬剤が、後に死亡率を増加させることが判明した⁸⁸⁾。より最近の事例としては、血糖値低下（重篤な糖尿病合併症、心血管イベント、死亡の代替指標）に基づいて承認された糖尿病治療薬（ロシグリタゾン）が、後に心不全による入院増加や心筋梗塞増加と関連していることが判明した事例が挙げられる⁸⁹⁾。また、BELLINI試験では、再発または難治性の多発性骨髄腫患者において無増悪生存期間（全生存期間の代替エンドポイン

ト）を改善した薬剤（ベネトクラックス）が、死亡率の上昇と関連していることが判明した⁹⁰⁾。

これらの有害事象は、代替エンドポイントや既知の疾患因果経路を介さない介入の意図していない結果など、様々な理由による可能性がある。また、代替エンドポイントが目標アウトカムと正の相関を示す対象者群においてさえ、介入が代替エンドポイントに好影響を与えない場合もある^{87,88)}。代替エンドポイントが主要アウトカムとして使用される場合、特に介入の潜在的な有害作用を明らかにし、代替エンドポイントに基づく結果を上書きする可能性がある場合には、目標アウトカムを副次的アウトカムとして収集することを推奨する。例えば、BELLINI試験では無増悪生存期間を主要アウトカムとしたが、全生存期間を副次的アウトカムとして収集したため、介入の有害性が捕捉され、試験が早期に中止された事例がある^{90,91)}。

CONSORT-Surrogate 項目 22.3 の例

[Examples of CONSORT-Surrogate item 22.3]

例 1: 追加解析の報告

[Example 1: Reporting subsequent analyses]

「追跡期間の中央値が約 17 か月であった時点で、全生存期間 [主要アウトカム] は不完全であった。しかしながら、KdD [カルフィルゾミブ、デキサメタゾン、ダラツムマブ] 群 (19%) では Kd 群 [カルフィルゾミブ、デキサメタゾン] (23%) と比較して死亡数が少なく、KdD 群と Kd 群の比較では全生存期間に正の有意傾向が認められた (付録 P7)。全生存期間は追跡調査の継続に伴い再評価される予定である⁹²⁾。」（角括弧内の文言を追加した。該当項目の報告時には使用を推奨する）

例 2: 22.1, 22.2 と 22.3 の統合項目（進行中の研究を報告）

[Example 2: Combining items 22.1, 22.2, and 22.3 (reports ongoing study)]

「本研究には限界があり、現時点では、ボソリチド治療が最終的な成人身長に及ぼす効果、およびこれが軟骨無形成症患者の機能性、生活の質、日常生活動作とどのように関連するかを直接評価することはできない。さらに、ボソリチド治療が軟骨無形成症に伴う医学的合併症を改善し、外科的介入の必要

性を減少させるかどうかは不明である。これらの制限に関する懸念は、低身長コミュニティの一部や支援団体にも共有されており、彼らは軟骨無形成症において身長増加（代替エンドポイント）のみを目的とする治療は優先すべきではなく、個人の短期的・長期的な健康状態（目標アウトカム）も同時に改善されるべきだと考えている。一方で、本試験参加者やその家族の一部からは、健康状態の改善も重要な成果であるものの、身長そのものの増加が環境へのアクセス向上、差別軽減、自尊心の向上につながるという見解も示されている。これらの限界、懸念、未解決の疑問に対処するため、進行中の非盲検第3相延長試験（ClinicalTrials.gov 番号 NCT03424018）では、本試験で報告された患者が最終的な成人身長に達するまで、ボソリチドの有益性と有害性のバランスを評価し続ける。本試験では、未治療の小人症児の登録データと比較し、生活の質、日常生活動作、医療・外科的介入の頻度と種類など、より広範な健康指標に関するボソリチド療法のデータを収集する。この長期研究は、軟骨無形成症の小児に対するボソリチド治療が、この疾患では欠如していると思われる思春期成長スパートをもたらすかどうかのデータを提供するとともに、長期治療に伴う有害事象を検出する機会を提供するものである⁹³⁾。（角括弧内の文言を追加した。該当項目の報告時には使用を推奨する。）

解説 [Explanation]

本項目は前項を踏まえ、目標アウトカムを用いて現在の知見（観察された有益性、有益性の欠如、有害性）を検証する後続の分析や研究を読者に伝えるものである。これには、介入が目標アウトカムに及ぼす効果を確認するための試験対象集団の長期追跡調査や、代替エンドポイント検証研究からのエビデンスが含まれる可能性がある。1990年から2011年の間に3つの高影響力ジャーナルに掲載された、代替エンドポイントを主要アウトカムとして使用した心血管試験の調査では、対象アウトカムを用いて結果を検証する後続試験があったのはわずか27%であったことが判明した⁹⁴⁾。がん領域では、米国食品医薬品局（FDA）による薬剤承認の遡及分析において、迅速承認の56%、従来型承認の37%が強力な代替エンドポイント検証エビデンスに裏付けら

れていなかったことが判明した⁹⁵⁾。それにもかかわらず、全生存期間という目標アウトカムに関する事後分析が行われた承認はわずか45%であった⁹⁵⁾。効果を検証する追跡研究の欠如は、心血管疾患・がん・薬剤関連介入を超えて広がり、有益性のない介入の継続的使用につながる可能性がある²³⁾。

今後の試験解析や追加研究の実施は、実現可能性や研究資金の確保状況など、複数の要因に依存することを認識している。さらに、こうした将来の分析研究の実施計画は時間の経過とともに変更される可能性がある。しかしながら、試験者にはこの項目について透明性をもって報告することを推奨する。すなわち、計画がないこと（その根拠を含む）を明示的に述べる、計画（研究期間を超えた追跡調査の計画、進行中の計画的または確認的目標アウトカム試験を含む）を記述する、あるいは当初の計画が変更されたことを記述することである。

その他の情報 [Other information]

新項目 [New items]

CONSORT-Surrogate 項目 26.1

[CONSORT-Surrogate extension item 26.1]

試験参加者が登録前に、試験が代替エンドポイントを用いて介入の効果を評価するように設計されていることを知らされたかどうか、またどのように知らされたかを明記

CONSORT-Surrogate 項目 26.2

[CONSORT-Surrogate extension item 26.2]

試験において代替エンドポイントおよび目標アウトカムのデータが収集された場合、将来の二次研究のためにデータのオープンアクセスの取り決めを明記

CONSORT-Surrogate 項目 26.1 の例

[Example of CONSORT-Surrogate item 26.1]

「参加者全員は、試験の性質、目的、考えられるリスク、利点 [代替エンドポイントを主要アウトカムとして使用した場合]、および IRB [機関審査委員会] によって承認されたインフォームド Consent プロトコルを使用した代替治療の選択肢について十分な情報を [受け取った]。参加者全員に、質問をしたり、試験への参加を検討したりする十分な時間と機会が [与えられた] ⁹⁶⁾。」（この例は項目を満

たしたものではないが、角括弧内の単語を使用して該当項目を報告する方法を示すために使用された。また、この例はプロトコルから取得され、過去形に変更された)

解説 [Explanation]

市民参画（地域参画とも呼ばれる）とは、研究活動やその成果を共有したり、関連する問題（倫理など）を議論したり、予備的な研究アイデアに関する意見を得るために、一般の人々の声に耳を傾け、交流し、つながることである⁹⁷⁾。患者と一般市民の関与は、特定の研究に焦点を当て、一般市民と共に、または一般市民によって行われる研究（一般市民の「ため」「対象」「について」ではなく）に関係する⁹⁷⁾。市民参画は、試験の計画と実施の両方に不可欠であるだけでなく、試験結果の解釈、そして試験参加者と一般市民のより大きな有益性にも重要である^{98,99)}。市民参画とインフォームドコンセントは、敬意を持って実施される研究の社会的価値を最大化するという同じ目標を持つ相互に支え合う問題である¹⁰⁰⁻¹⁰²⁾。

インフォームドコンセントは、研究に参加する前に参加者となる人間を対象とする研究の法的および倫理的要件である^{103,104)}。参加者に、参加によって期待される有益性や潜在的なリスクを含む試験の詳細を適切に伝えることが含まれる^{103,105)}。したがって、代替エンドポイントを主要アウトカムとして用いる試験においては、インフォームドコンセントは、試験参加者に代替エンドポイントの使用とその有害性と有益性について理解を深めてもらうための手段となり、理想的には継続的な理解を深めてもらうための手段となる。しかしながら、初期段階の試験（その多くは代替エンドポイントに依存している可能性がある）から得られたエビデンスは、参加者への有害性と有益性に関するコミュニケーションが最適ではないことを示唆している。172件の初期段階の試験のインフォームドコンセント文書を調査したところ、言及された健康上の有益性のアウトカムを明記したのはわずか45%であり、健康リスクの可能性を明記したのはわずか63%で、そのうちリスクが研究手順によるものか、潜在的に有益な介入によるものかが具体的に記載されていたのは半数のみであった¹⁰⁵⁾。

参加者に、研究で代替エンドポイント（および関連する限界）が使用されたことを伝えることは、インフォームドコンセントにとって非常に重要である。この項目については、SPIRIT-Surrogateにおいて詳細に議論し、その実施方法に関する提案も示している。本項目を実施しない試験については、著者がその正当性を説明すべきである。

CONSORT-Surrogate 項目 26.2 の例

[Examples of CONSORT-Surrogate item 26.2]

例 1: データはリクエストに応じて利用可能 [Example 1: Data available on request]

「データ共有に関する声明: 完全に匿名化された患者データセットは、データ利用が承認された研究者に公開される（リクエストの宛先は ctu.beatlupus@ucl.ac.uk である）¹⁰⁷⁾。」（試験の主要アウトカムは血清 IgG 中の抗二本鎖 DNA 抗体レベルの代替エンドポイントであり、疾患の再燃は副次的アウトカムである目標アウトカムであった）

例 2: 仲介者を通じて利用できるデータ

[Example 2: Data available via an intermediary]

「ヤンセンは、医学の知識と公衆衛生の向上につながる科学研究のために、研究者や医師からの臨床研究報告書や参加者レベルのデータの要請を評価する独立審査委員会として機能するため、Yale Open Data Access (YODA) プロジェクトと契約を結んでいる。データは、YODA による正式なリクエストと明確な分析計画の承認を得て公開され次第、公開される。このプロセスの詳細、またはリクエストについては、YODA プロジェクトのウェブサイト (<https://yoda.yale.edu>) を参照されたい。ジョンソン・エンド・ジョンソンのヤンセン製薬会社のデータ共有ポリシーは、<https://www.janssen.com/clinical-trials/transparency> に掲載されている¹⁰⁸⁾。」（この第1相試験は安全性を主要アウトカムとしていましたが、中間機関へのデータの預託の例でもある）

解説 [Explanation]

代替エンドポイントを主要アウトカムとして使用する場合における目標アウトカムデータを収集することの重要性については既に解説している。すなわち、代替エンドポイントの妥当性を検証し、介入に

よる有害事象のモニタリングに貢献するためである。したがって、試験チームには、目標とするアウトカムを副次的アウトカムとして収集することを推奨している。代替指標の妥当性検証研究を実施する上で重要な課題は、完了した研究から個々の参加者データへのアクセスが限られていることである¹⁰⁹⁾。したがって、代替エンドポイントと目標アウトカムのデータが収集された場合、それらを共有することで、治験データセットを活用して代替エンドポイントの検証分野を発展させることができる。

この項目について適切な報告（つまり、データが利用可能であるという記述）を行うだけでは不十分である。試験責任医師チームは、データセットを共有することに真摯に取り組むべきである。最近発表された臨床試験に関する調査によると、ほとんどの試験著者がデータ共有の意向を表明しているにもかかわらず、個々の患者レベルのデータへのアクセスは圧倒的に低い（25%未満）ことが判明した。データ共有には、参加者の機密性喪失のリスク、データの不適切な利用に対する懸念、データへのアクセス権を持つ同分野の研究者との競争など、様々な課題が存在する可能性がある^{110,111)}。したがって、データ共有が不可能な場合、あるいはデータの一部のみ共有可能な場合は、試験著者はその旨を明記し、正当な理由を示すべきである。

結論 [Conclusions]

代替エンドポイントを使用する試験では、より

適切で透明性の高い報告が求められる。CONSORT-Surrogateは、代替エンドポイントを主要アウトカムとして用いた試験報告および出版物に対する最小限の報告要件を示すものである。CONSORT-Surrogateを適切に適用することにより、この種の試験の報告の質が向上し、結果の解釈が容易になるとともに、新たな研究の実施や政策決定に資することが期待される。なお、本拡張ガイドラインは、CONSORT 2010の主報告ガイドラインと併せて使用する必要がある。

CONSORT-Surrogateは、研究の無駄の削減に貢献し得るものである³⁾。しかしながら、多くのジャーナルが主要なCONSORTチェックリストの使用を推奨している一方で、その拡張版の使用を推奨しているジャーナルはほとんどない¹⁾。もっとも、この拡張版に従うこと自体が、リサーチクエストの不適切な設定、バイアス、不適切な研究デザインなど、研究の無駄を生むその他の要因を排除するものではない²⁾。特に、試験チームおよび読者は、代替エンドポイント測定における偏りが、目標アウトカムに対する介入効果の予測精度の低下につながる可能性があることに留意する必要がある¹⁶⁾。

最後に、本拡張ガイドラインの全ての項目が適切に報告されたとしても、それが試験研究チームやより広範な科学コミュニティに対して、可能な限り目標アウトカムに対する介入効果を直接評価し、報告する努力を妨げるものではない。

参考文献 [REFERENCES]

1. Junqueira DR, Zorzela L, Golder S, *et al.* CONSORT Harms Group. CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *BMJ* 2023;**381**: e073725. doi:10.1136/bmj-2022-073725. pmid:37094878
2. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; **374**: 86-9. doi:10.1016/S0140-6736(09)60329-9. pmid:19525005CrossRefPubMedWeb of ScienceGoogle Scholar
3. Glasziou P, Altman DG, Bossuyt P, *et al.* Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014; **383**: 267-76. doi:10.1016/S0140-6736(13)62228-X. pmid:24411647CrossRefPubMedWeb of ScienceGoogle Scholar
4. 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. pmid:20332511FREE Full TextGoogle Scholar
5. Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev* 2012; **1**: 60. doi:10.1186/2046-4053-1-60. pmid:23194585CrossRefPubMedGoogle Scholar
6. 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**: 814-22. doi:10.1001/jama.2013.879.

- pmid:23443445CrossRefPubMedWeb of ScienceGoogle Scholar
7. Butcher NJ, Monsour A, Mew EJ, *et al.* Guidelines for Reporting Outcomes in Trial Reports: The CONSORT-Outcomes 2022 Extension. *JAMA* 2022; **328**: 2252-64. doi:10.1001/jama.2022.21022. pmid:36511921CrossRefPubMedGoogle Scholar
 8. FDA-NIH Biomarker Working Group. BEST (biomarkers, endpoints, and other tools) resource. 2016. https://www.ncbi.nlm.nih.gov/books/NBK326791/pdf/Bookshelf_NBK326791.pdf.
 9. EUnetHTA. EUnetHTA 21 – Individual Practical Guideline Document D4.4 – OUTCOMES (ENDPOINTS), 2023.
 10. Azizi M, Sanghvi K, Saxena M, *et al.* RADIANCE-HTN investigators. Ultrasound renal denervation for hypertension resistant to a triple medication pill (RADIANCE-HTN TRIO): a randomised, multicentre, single-blind, sham-controlled trial. *Lancet* 2021; **397**: 2476-86. doi:10.1016/S0140-6736(21)00788-1. pmid:34010611CrossRefPubMedGoogle Scholar
 11. Skoulidis F, Li BT, Dy GK, *et al.* Sotorasib for Lung Cancers with KRAS p.G12C Mutation. *N Engl J Med* 2021; **384**: 2371-81. doi:10.1056/NEJMoa2103695. pmid:34096690CrossRefPubMedGoogle Scholar
 12. Wright ME, Delacroix E, Sonnevile KR, *et al.* Reducing paediatric overweight and obesity through motivational interviewing: study protocol for a randomised controlled trial in the AAP PROS research network. *BMJ Open* 2020; **10**: e035720. doi:10.1136/bmjopen-2019-035720. pmid:32723736Abstract/FREE Full TextGoogle Scholar
 13. Chan A-W, Tetzlaff JM, Gøtzsche PC, *et al.* SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013; **346**: e7586. doi:10.1136/bmj.e7586. pmid:23303884Abstract/FREE Full TextGoogle Scholar
 14. 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. pmid:28529187CrossRefPubMedGoogle Scholar
 15. 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**: 852-60. doi:10.1056/NEJMs1012065. pmid:21366476CrossRefPubMedWeb of ScienceGoogle Scholar
 16. Manyara AM, Davies P, Stewart D, *et al.* Definitions, acceptability, limitations, and guidance in the use and reporting of surrogate end points in trials: a scoping review. *J Clin Epidemiol* 2023; **160**: 83-99. doi:10.1016/j.jclinepi.2023.06.013. pmid:37380118CrossRefPubMedGoogle Scholar
 17. Patel RB, Vaduganathan M, Samman-Tahhan A, *et al.* Trends in Utilization of Surrogate Endpoints in Contemporary Cardiovascular Clinical Trials. *Am J Cardiol* 2016; **117**: 1845-50. doi:10.1016/j.amjcard.2016.03.021. pmid:27085935CrossRefPubMedGoogle Scholar
 18. Ciani O, Buyse M, Garside R, *et al.* Meta-analyses of randomized controlled trials show suboptimal validity of surrogate outcomes for overall survival in advanced colorectal cancer. *J Clin Epidemiol* 2015; **68**: 833-42. doi:10.1016/j.jclinepi.2015.02.016. pmid:25863582CrossRefPubMedGoogle Scholar
 19. Gandhi GY, Murad MH, Fujiyoshi A, *et al.* Patient-important outcomes in registered diabetes trials. *JAMA* 2008; **299**: 2543-9. doi:10.1001/jama.299.21.2543. pmid:18523223CrossRefPubMedWeb of ScienceGoogle Scholar
 20. la Cour JL, Brok J, Gøtzsche PC. Inconsistent reporting of surrogate outcomes in randomised clinical trials: cohort study. *BMJ* 2010; **341**: c3653. doi:10.1136/bmj.c3653. pmid:20719823Abstract/FREE Full TextGoogle Scholar
 21. Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. *Value Health* 2017; **20**: 487-95. doi:10.1016/j.jval.2016.10.011. pmid:28292495CrossRefPubMedGoogle Scholar
 22. Ciani O, Buyse M, Garside R, *et al.* Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study. *BMJ* 2013; **346**: f457. doi:10.1136/bmj.f457. pmid:23360719Abstract/FREE Full TextGoogle Scholar
 23. Dawoud D, Naci H, Ciani O, Bujkiewicz S. Raising the bar for using surrogate endpoints in drug regulation and health technology assessment. *BMJ* 2021; **374**: n2191. doi:10.1136/bmj.n2191. pmid:34526320FREE Full TextGoogle Scholar
 24. Manyara AM, Davies P, Stewart D, *et al.* Reporting of surrogate endpoints in randomised controlled trial protocols (SPIRIT-Surrogate): extension checklist with explanation and elaboration. *BMJ* 2024; **386**: e078525. doi:10.1136/bmj-2023-078525 pmid:33507252FREE Full TextGoogle Scholar
 25. Ferreira-González I, Permanyer-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**: 651-7, discussion 658-62. doi:10.1016/j.jclinepi.2006.10.020. pmid:17573977CrossRefPubMedGoogle Scholar
 26. Butcher NJ, Mew EJ, Monsour A, Chan AW, Moher D, Offringa M. Outcome reporting recommendations for clinical trial protocols and reports: a scoping review. *Trials* 2020; **21**: 620. doi:10.1186/s13063-020-04440-w pmid:32641085CrossRefPubMedGoogle Scholar
 27. Ciani O, Manyara AM, Davies P, *et al.* A framework for the definition and interpretation of the use of surrogate endpoints in interventional trials. *E Clinical Medicine* 2023; **65**: 102283. doi:10.1016/j.eclinm.2023.102283 pmid:37877001CrossRefPubMedGoogle Scholar
 28. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010; **7**: e1000217. doi:10.1371/journal.pmed.1000217. pmid:20169112CrossRefPubMedGoogle Scholar
 29. EQUATOR. CONSORT-SURROGATE – CONSORT extension for trials using surrogate primary endpoints 2022. <https://www>.

equator-network.org/library/reporting-guidelines-under-development/reporting-guidelines-under-development-for-clinical-trials/#SURROGATE.

30. Manyara AM, Davies P, Stewart D, *et al.* Protocol for the development of SPIRIT and CONSORT extensions for randomised controlled trials with surrogate primary endpoints: SPIRIT-SURROGATE and CONSORT-SURROGATE. *BMJ Open* 2022; **12**: e064304. doi:10.1136/bmjopen-2022-064304. pmid:36220321CrossRefPubMedGoogle Scholar
31. Manyara AM, Davies P, Stewart D, *et al.* Scoping and targeted reviews to support development of SPIRIT and CONSORT extensions for randomised controlled trials with surrogate primary endpoints: protocol. *BMJ Open* 2022; **12**: e062798. doi:10.1136/bmjopen-2022-062798. pmid:36229145Abstract/FREE Full TextGoogle Scholar
32. Ciani O, Manyara A, Taylor RS. Need for better reporting of trials with surrogate endpoints: SPIRIT/CONSORT-SURROGATE extensions. *J Epidemiol Community Health* 2022; **76**: 769-70. doi:10.1136/jech-2022-219294. pmid:35750481FREE Full TextGoogle Scholar
33. Ciani O, Manyara AM, Chan A-W, Taylor RS, SPIRIT-SURROGATE/CONSORT-SURROGATE project group. Surrogate endpoints in trials: a call for better reporting. *Trials* 2022; **23**: 991. doi:10.1186/s13063-022-06904-7. pmid:36503559CrossRefPubMedGoogle Scholar
34. Ciani O, Manyara AM, Taylor RS. Surrogate endpoints in trials-a call for better reporting. *BMJ* 2022; **378**: o1912. doi:10.1136/bmj.o1912. pmid:35905986FREE Full TextGoogle Scholar
35. Ciani O, Manyara AM, Taylor RS. Surrogate end points in cardio-thoracic trials: a call for better reporting and improved interpretation of trial findings. *Eur J Cardiothorac Surg* 2022; **62**: ezac449. doi:10.1093/ejcts/ezac449. pmid:36112148CrossRefPubMedGoogle Scholar
36. Manyara AM, Ciani O, Taylor RS. A call for better reporting of trials using surrogate primary endpoints. *Alzheimers Dement (N Y)* 2022; **8**: e12340. doi:10.1002/trc2.12340. pmid:35910671CrossRefPubMedGoogle Scholar
37. Hopewell S, Clarke M, Moher D, *et al.* CONSORT Group. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 2008; **5**: e20. doi:10.1371/journal.pmed.0050020. pmid:18215107CrossRefPubMedGoogle Scholar
38. Panagiotou A, Trendelenburg M, Heijnen IAFM, *et al.* A Randomized Trial of Recombinant Human C1-Esterase-Inhibitor in the Prevention of Contrast-Induced Kidney Injury. *JACC Cardiovasc Interv* 2020; **13**: 833-42. doi:10.1016/j.jcin.2019.11.021. pmid:32171721Abstract/FREE Full TextGoogle Scholar
39. de Boer SA, Heerspink HJL, Juárez Orozco LE, *et al.* Effect of linagliptin on pulse wave velocity in early type 2 diabetes: A randomized, double-blind, controlled 26-week trial (RELEASE). *Diabetes Obes Metab* 2017; **19**: 1147-54. doi:10.1111/dom.12925. pmid:28244635CrossRefPubMedGoogle Scholar
40. Bandholm T, Thorborg K, Ardern CL, Christensen R, Henriksen M. Writing up your clinical trial report for a scientific journal: the REPORT trial guide for effective and transparent research reporting without spin. *Br J Sports Med* 2022; **56**: 683-91. doi:10.1136/bjsports-2021-105058. pmid:35193854Abstract/FREE Full TextGoogle Scholar
41. Bauchner H, Golub RM, Fontanarosa PB. Reporting and Interpretation of Randomized Clinical Trials. *JAMA* 2019; **322**: 732-5. doi:10.1001/jama.2019.12056. pmid:31454026CrossRefPubMedGoogle Scholar
42. Howard BV, Roman MJ, Devereux RB, *et al.* Effect of lower targets for blood pressure and LDL cholesterol on atherosclerosis in diabetes: the SANDS randomized trial. *JAMA* 2008; **299**: 1678-89. doi:10.1001/jama.299.14.1678. pmid:18398080CrossRefPubMedWeb of ScienceGoogle Scholar
43. Schuster Bruce C, Brhlikova P, Heath J, McGettigan P. The use of validated and nonvalidated surrogate endpoints in two European Medicines Agency expedited approval pathways: A cross-sectional study of products authorised 2011-2018. *PLoS Med* 2019; **16**: e1002873. doi:10.1371/journal.pmed.1002873 pmid:31504034CrossRefPubMedGoogle Scholar
44. Hey SP, Kesselheim AS, Patel P, Mehrotra P, Powers JH 3rd.. US Food and Drug Administration Recommendations on the Use of Surrogate Measures as End Points in New Anti-infective Drug Approvals. *JAMA Intern Med* 2020; **180**: 131-8. doi:10.1001/jamainternmed.2019.5451. pmid:31710344CrossRefPubMedGoogle Scholar
45. Block GA, Bushinsky DA, Cheng S, *et al.* Effect of Etelcalcetide vs Cinacalcet on Serum Parathyroid Hormone in Patients Receiving Hemodialysis With Secondary Hyperparathyroidism: A Randomized Clinical Trial. *JAMA* 2017; **317**: 156-64. doi:10.1001/jama.2016.19468. pmid:28097356CrossRefPubMedGoogle Scholar
46. Etehad D, Emdin CA, Kiran A, *et al.* Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. *Lancet* 2016; **387**: 957-67. doi:10.1016/S0140-6736(15)01225-8. pmid:26724178CrossRefPubMedGoogle Scholar
47. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics* 2000; **1**: 49-67. doi:10.1093/biostatistics/1.1.49. pmid:12933525CrossRefPubMedGoogle Scholar
48. Buyse M, Saad ED, Burzykowski T, Regan MM, Sweeney CS. Surrogacy Beyond Prognosis: The Importance of “Trial-Level” Surrogacy. *Oncologist* 2022; **27**: 266-71. doi:10.1093/oncolo/oyac006. pmid:35380717CrossRefPubMedGoogle Scholar
49. Alonso A, Van der Elst W, Molenberghs G, Buyse M, Burzykowski T. On the relationship between the causal-inference and meta-analytic paradigms for the validation of surrogate endpoints. *Biometrics* 2015; **71**: 15-24. doi:10.1111/biom.12245. pmid:25274284CrossRefPubMedGoogle Scholar
50. Alonso A, Van der Elst W, Molenberghs G, Buyse M, Burzykowski T. An information-theoretic approach for the evaluation of

- surrogate endpoints based on causal inference. *Biometrics* 2016; **72**: 669-77. doi:10.1111/biom.12483. pmid:26864244CrossRefPubMedGoogle Scholar
51. Bujkiewicz S, Jackson D, Thompson JR, *et al*. Bivariate network meta-analysis for surrogate endpoint evaluation. *Stat Med* 2019; **38**: 3322-41. doi:10.1002/sim.8187 pmid:31131475CrossRefPubMedGoogle Scholar
 52. Papanikos T, Thompson JR, Abrams KR, *et al*. Bayesian hierarchical meta-analytic methods for modeling surrogate relationships that vary across treatment classes using aggregate data. *Stat Med* 2020; **39**: 1103-24. doi:10.1002/sim.8465 pmid:31990083CrossRefPubMedGoogle Scholar
 53. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989; **8**: 431-40. doi:10.1002/sim.4780080407. pmid:2727467CrossRefPubMedWeb of ScienceGoogle Scholar
 54. Weir CJ, Taylor RS. Informed decision-making: Statistical methodology for surrogacy evaluation and its role in licensing and reimbursement assessments. *Pharm Stat* 2022; **21**: 740-56. doi:10.1002/pst.2219. pmid:35819121CrossRefPubMedGoogle Scholar
 55. Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Stat Med* 2006; **25**: 183-203. doi:10.1002/sim.2319. pmid:16252272CrossRefPubMedWeb of ScienceGoogle Scholar
 56. Molenberghs G, Burzykowski T, Alonso A, *et al*. The meta-analytic framework for the evaluation of surrogate endpoints in clinical trials. *J Stat Plan Inference* 2008; **138**: 432-49. doi:10.1016/j.jspi.2007.06.005.CrossRefGoogle Scholar
 57. Institute for Quality and Efficiency in Health Care (IQWiG). Validity of surrogate endpoints in oncology: Executive summary of rapid report A10-05, Version 1.1. Institute for Quality and Efficiency in Health Care: Executive Summaries. 2005.
 58. Lassere MN, Johnson KR, Schiff M, Rees D. Is blood pressure reduction a valid surrogate endpoint for stroke prevention? An analysis incorporating a systematic review of randomised controlled trials, a by-trial weighted errors-in-variables regression, the surrogate threshold effect (STE) and the Biomarker-Surrogacy (BioSurrogate) Evaluation Schema (BSES). *BMC Med Res Methodol* 2012; **12**: 27. doi:10.1186/1471-2288-12-27. pmid:22409774CrossRefPubMedGoogle Scholar
 59. Boissel JP, Collet JP, Moleur P, Haugh M. Surrogate endpoints: a basis for a rational approach. *Eur J Clin Pharmacol* 1992; **43**: 235-44. doi:10.1007/BF02333016. pmid:1425885CrossRefPubMedWeb of ScienceGoogle Scholar
 60. Xie W, Halabi S, Tierney JF, *et al*. A Systematic Review and Recommendation for Reporting of Surrogate Endpoint Evaluation Using Meta-analyses. *JNCI Cancer Spectr* 2019; **3**: pkz002. doi:10.1093/jncics/pkz002. pmid:31360890CrossRefPubMedGoogle Scholar
 61. la Cour JL, Brok J, Gøtzsche PC. Inconsistent reporting of surrogate outcomes in randomised clinical trials: cohort study. *BMJ* 2010; **341**: c3653. doi:10.1136/bmj.c3653. pmid:20719823Abstract/FREE Full TextGoogle Scholar
 62. Buyse M, Molenberghs G, Paoletti X, *et al*. Statistical evaluation of surrogate endpoints with examples from cancer clinical trials. *Biom J* 2016; **58**: 104-32. doi:10.1002/bimj.201400049. pmid:25682941CrossRefPubMedGoogle Scholar
 63. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. *Stat Med* 1997; **16**: 1965-82. doi:10.1002/(SICI)1097-0258(19970915)16:17<1965::AID-SIM630>3.0.CO;2-M. pmid:9304767CrossRefPubMedWeb of ScienceGoogle Scholar
 64. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics* 2002; **58**: 21-9. doi:10.1111/j.0006-341X.2002.00021.x. pmid:11890317CrossRefPubMedWeb of ScienceGoogle Scholar
 65. Tibaldi F, Abrahantes JC, Molenberghs G, *et al*. Simplified hierarchical linear models for the evaluation of surrogate endpoints. *J Stat Comput Simul* 2003; **73**: 643-58. doi:10.1080/0094965031000062177.CrossRefGoogle Scholar
 66. Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Stat* 2006; **5**: 173-86. doi:10.1002/pst.207. pmid:17080751CrossRefPubMedWeb of ScienceGoogle Scholar
 67. Alonso A, Molenberghs G. Surrogate marker evaluation from an information theory perspective. *Biometrics* 2007; **63**: 180-6. doi:10.1111/j.1541-0420.2006.00634.x. pmid:17447943CrossRefPubMedWeb of ScienceGoogle Scholar
 68. Belin L, Tan A, De Rycke Y, Dechartres A. Progression-free survival as a surrogate for overall survival in oncology trials: a methodological systematic review. *Br J Cancer* 2020; **122**: 1707-14. doi:10.1038/s41416-020-0805-y. pmid:32214230CrossRefPubMedGoogle Scholar
 69. Haase CL, Lopes S, Olsen AH, Satyrganova A, Schneck V, McEwan P. Weight loss and risk reduction of obesity-related outcomes in 0.5 million people: evidence from a UK primary care database. *Int J Obes (Lond)* 2021; **45**: 1249-58. doi:10.1038/s41366-021-00788-4. pmid:33658682CrossRefPubMedGoogle Scholar
 70. Winter JE, MacInnis RJ, Wattanapenpaiboon N, Nowson CA. BMI and all-cause mortality in older adults: a meta-analysis. *Am J Clin Nutr* 2014; **99**: 875-90. doi:10.3945/ajcn.113.068122. pmid:24452240Abstract/FREE Full TextGoogle Scholar
 71. Shah SJ, Voors AA, McMurray JJV, *et al*. Effect of Neladenoson Bialanate on Exercise Capacity Among Patients With Heart Failure With Preserved Ejection Fraction: A Randomized Clinical Trial. *JAMA* 2019; **321**: 2101-12. doi:10.1001/jama.2019.6717. pmid:31162568CrossRefPubMedGoogle Scholar
 72. 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. doi:10.1136/bmj.k3750. pmid:30560792FREE Full TextGoogle Scholar
 73. Tasanarong A, Hutayanon P, Piyayotai D. Urinary Neutrophil Gelatinase-Associated Lipocalin predicts the severity of contrast-induced acute kidney injury in chronic kidney disease patients undergoing elective coronary procedures. *BMC Nephrol* 2013; **14**: 270. doi:10.1186/1471-2369-14-270. pmid:24305547CrossRefPubMedGoogle Scholar
 74. van de Putte LB, Atkins C, Malaise M, *et al*. Efficacy and safety of adalimumab as monotherapy in patients with rheumatoid arthritis

- for whom previous disease modifying antirheumatic drug treatment has failed. *Ann Rheum Dis* 2004; **63**: 508-16. doi:10.1136/ard.2003.013052. pmid:15082480Abstract/FREE Full TextGoogle Scholar
75. Parekh DJ, Reis IM, Castle EP, *et al.* Robot-assisted radical cystectomy versus open radical cystectomy in patients with bladder cancer (RAZOR): an open-label, randomised, phase 3, non-inferiority trial. *Lancet* 2018; **391**: 2525-36. doi:10.1016/S0140-6736(18)30996-6. pmid:29976469CrossRefPubMedGoogle Scholar
76. 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. pmid:20719825Abstract/FREE Full TextGoogle Scholar
77. Lim E, Brown A, Helmy A, Mussa S, Altman DG. Composite outcomes in cardiovascular research: a survey of randomized trials. *Ann Intern Med* 2008; **149**: 612-7. doi:10.7326/0003-4819-149-9-200811040-00004. pmid:18981486CrossRefPubMedWeb of ScienceGoogle Scholar
78. Wells GA, Tugwell P, Tomasson G, *et al.* Composite outcomes at OMERACT: Multi-outcome domains and composite outcome domains. *Semin Arthritis Rheum* 2021; **51**: 1370-7. doi:10.1016/j.semarthrit.2021.11.001. pmid:34863558CrossRefPubMedGoogle Scholar
79. Hochman M, McCormick D. Endpoint selection and relative (versus absolute) risk reporting in published medication trials. *J Gen Intern Med* 2011; **26**: 1246-52. doi:10.1007/s11606-011-1813-7. pmid:21842324CrossRefPubMedGoogle Scholar
80. Delgado A, Guddati AK. Clinical endpoints in oncology - a primer. *Am J Cancer Res* 2021; **11**: 1121-31. pmid:33948349PubMedGoogle Scholar
81. Sorbello A, Komo S, Valappil T, Nambiar S. Registration trials of antibacterial drugs for the treatment of nosocomial pneumonia. *Clin Infect Dis* 2010; **51**(Suppl 1): S36-41. doi:10.1086/653038. pmid:20597669CrossRefPubMedWeb of ScienceGoogle Scholar
82. Timsit JF, de Kraker MEA, Sommer H, *et al.* COMBACTE-NET consortium. Appropriate endpoints for evaluation of new antibiotic therapies for severe infections: a perspective from COMBACTE's STAT-Net. *Intensive Care Med* 2017; **43**: 1002-12. doi:10.1007/s00134-017-4802-4. pmid:28466147CrossRefPubMedGoogle Scholar
83. Heindel W, Weigel S, Gerß J, *et al.* TOSYMA Screening Trial Study Group. Digital breast tomosynthesis plus synthesised mammography versus digital screening mammography for the detection of invasive breast cancer (TOSYMA): a multicentre, open-label, randomised, controlled, superiority trial. *Lancet Oncol* 2022; **23**: 601-11. doi:10.1016/S1470-2045(22)00194-2. pmid:35427470CrossRefPubMedGoogle Scholar
84. Baechele C, Scherler W, Lang A, Filla T, Kuss O. Is HbA1c a valid surrogate for mortality in type 2 diabetes? Evidence from a meta-analysis of randomized trials. *Acta Diabetol* 2022; **59**: 1257-63. doi:10.1007/s00592-022-01887-y. pmid:35534726CrossRefPubMedGoogle Scholar
85. Ciani O, Piepoli M, Smart N, *et al.* Validation of Exercise Capacity as a Surrogate Endpoint in Exercise-Based Rehabilitation for Heart Failure: A Meta-Analysis of Randomized Controlled Trials. *JACC Heart Fail* 2018; **6**: 596-604. doi:10.1016/j.jchf.2018.03.017. pmid:29957192Abstract/FREE Full TextGoogle Scholar
86. Green E, Yothers G, Sargent DJ. Surrogate endpoint validation: statistical elegance versus clinical relevance. *Stat Methods Med Res* 2008; **17**: 477-86. doi:10.1177/0962280207081863. pmid:18285438CrossRefPubMedGoogle Scholar
87. Vanderweele TJ. Surrogate measures and consistent surrogates. *Biometrics* 2013; **69**: 561-9. doi:10.1111/biom.12071. pmid:24073861CrossRefPubMedWeb of ScienceGoogle Scholar
88. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996; **125**: 605-13. doi:10.7326/0003-4819-125-7-199610010-00011. pmid:8815760CrossRefPubMedWeb of ScienceGoogle Scholar
89. Cohen D. Rosiglitazone: what went wrong? *BMJ* 2010; **341**: c4848. doi:10.1136/bmj.c4848 pmid:20819889FREE Full TextGoogle Scholar
90. Kumar S, Rajkumar SV. Surrogate endpoints in randomised controlled trials: a reality check. *Lancet* 2019; **394**: 281-3. doi:10.1016/S0140-6736(19)31711-8. pmid:31354129CrossRefPubMedGoogle Scholar
91. Kumar SK, Harrison SJ, Cavo M, *et al.* Venetoclax or placebo in combination with bortezomib and dexamethasone in patients with relapsed or refractory multiple myeloma (BELLINI): a randomised, double-blind, multicentre, phase 3 trial. *Lancet Oncol* 2020; **21**: 1630-42. doi:10.1016/S1470-2045(20)30525-8. pmid:33129376CrossRefPubMedGoogle Scholar
92. Dimopoulos M, Quach H, Mateos M-V, *et al.* Carfilzomib, dexamethasone, and daratumumab versus carfilzomib and dexamethasone for patients with relapsed or refractory multiple myeloma (CANDOR): results from a randomised, multicentre, open-label, phase 3 study. *Lancet* 2020; **396**: 186-97. doi:10.1016/S0140-6736(20)30734-0. pmid:32682484CrossRefPubMedGoogle Scholar
93. Savarirayan R, Tofts L, Irving M, *et al.* Once-daily, subcutaneous vosoritide therapy in children with achondroplasia: a randomised, double-blind, phase 3, placebo-controlled, multicentre trial. *Lancet* 2020; **396**: 684-92. doi:10.1016/S0140-6736(20)31541-5. pmid:32891212CrossRefPubMedGoogle Scholar
94. Bikdeli B, Punnathinon N, Akram Y, *et al.* Two Decades of Cardiovascular Trials With Primary Surrogate Endpoints: 1990-2011. *J Am Heart Assoc* 2017; **6**: e005285. doi:10.1161/JAHA.116.005285. pmid:28325713Abstract/FREE Full TextGoogle Scholar
95. Kim C, Prasad V. Strength of Validation for Surrogate End Points Used in the US Food and Drug Administration's Approval of Oncology Drugs. *Mayo Clin Proc* 2016; **10**: S0025-6196(16)00125-7. doi:10.1016/j.mayocp.2016.02.012. pmid:27236424CrossRefPubMedGoogle Scholar
96. Koshizaka M, Ishikawa K, Ishikawa T, *et al.* PRIME-V Study Investigators. Efficacy and safety of ipragliflozin and metformin for

- visceral fat reduction in patients with type 2 diabetes receiving treatment with dipeptidyl peptidase-4 inhibitors in Japan: a study protocol for a prospective, multicentre, blinded-endpoint phase IV randomised controlled trial (PRIME-V study). *BMJ Open* 2017; **7**: e015766. doi:10.1136/bmjopen-2016-015766. pmid:28490565Abstract/FREE Full TextGoogle Scholar
97. NIHR School for Primary Care Research. What is Patient and Public Involvement and Public Engagement? 2023. <https://www.spcr.nihr.ac.uk/PPI>.
 98. Geißler J, Isham E, Hickey G, Ballard C, Corbett A, Lubbert C. Patient involvement in clinical trials. *Commun Med (Lond)*2022; **2**: 94. doi:10.1038/s43856-022-00156-x. pmid:35903184CrossRefPubMedGoogle Scholar
 99. Selman LE, Clement C, Douglas M, *et al*. Patient and public involvement in randomised clinical trials: a mixed-methods study of a clinical trials unit to identify good practice, barriers and facilitators. *Trials* 2021; **22**: 735. doi:10.1186/s13063-021-05701-y. pmid:34688304CrossRefPubMedGoogle Scholar
 100. Davies A, Ormel I, Bernier A, *et al*. A rapid review of community engagement and informed consent processes for adaptive platform trials and alternative design trials for public health emergencies [version 1; peer review: 2 approved]. *Wellcome Open Res* 2023; **8**: 194. doi:10.12688/wellcomeopenres.19318.1. pmid:37654739CrossRefPubMedGoogle Scholar
 101. Molyneux S, Bull S. Consent and Community Engagement in Diverse Research Contexts: Reviewing and Developing Research and Practice: Participants in the Community Engagement and Consent Workshop, Kilifi, Kenya, March 2011. *J Empir Res Hum Res Ethics* 2013; **8**: 1-18. doi:10.1525/jer.2013.8.4.1.CrossRefPubMedGoogle Scholar
 102. Zulu JM, Sandøy IF, Moland KM, Musonda P, Munsaka E, Blystad A. The challenge of community engagement and informed consent in rural Zambia: an example from a pilot study. *BMC Med Ethics* 2019; **20**: 45. doi:10.1186/s12910-019-0382-x. pmid:31272489CrossRefPubMedGoogle Scholar
 103. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013; **310**: 2191-4. doi:10.1001/jama.2013.281053. pmid:24141714CrossRefPubMedWeb of ScienceGoogle Scholar
 104. Council for International Organizations of Medical Sciences. International ethical guidelines for biomedical research involving human subjects. *Bull Med Ethics* 2002; (182):17-23.pmid:14983848PubMedGoogle Scholar
 105. Kahrs H, Bossert S, Schürmann C, Strech D. Details of risk-benefit communication in informed consent documents for phase I/II trials. *Clin Trials* 2021; **18**: 71-80. doi:10.1177/1740774520971770. pmid:33231107CrossRefPubMedGoogle Scholar
 106. Daly TP. Informing consent to antibodies in Alzheimer’s disease. *BMJ* 2023; **383**: 2350. doi:10.1136/bmj.p2350. pmid:37821123CrossRefPubMedGoogle Scholar
 107. Shipa M, Embleton-Thirsk A, Parvaz M, *et al*. BEAT-LUPUS Investigators. Effectiveness of Belimumab After Rituximab in Systemic Lupus Erythematosus : A Randomized Controlled Trial. *Ann Intern Med* 2021; **174**: 1647-57. doi:10.7326/M21-2078. pmid:34698499CrossRefPubMedGoogle Scholar
 108. Bockstal V, Shukarev G, McLean C, *et al*. First-in-human study to evaluate safety, tolerability, and immunogenicity of heterologous regimens using the multivalent filovirus vaccines Ad26.Filo and MVA-BN-Filo administered in different sequences and schedules: A randomized, controlled study. *PLoS One* 2022; **17**: e0274906. doi:10.1371/journal.pone.0274906. pmid:36197845CrossRefPubMedGoogle Scholar
 109. Taylor R, Ciani O. Response to Wang *et al*. Quality of individual participant data (IPD) meta-analyses reporting might need improving but leveraging access to IPD is a more fundamental problem. *BMJ* 2021. Accessed 5 November 2023. <https://www.bmj.com/content/373/bmj.n736/tr-0>Google Scholar
 110. Danchev V, Min Y, Borghi J, Baiocchi M, Ioannidis JPA. Evaluation of Data Sharing After Implementation of the International Committee of Medical Journal Editors Data Sharing Statement Requirement. *JAMA Netw Open*2021; **4**: e2033972. doi:10.1001/jamanetworkopen.2020.33972. pmid:33507256CrossRefPubMedGoogle Scholar
 111. Esmail LC, Kapp P, Assi R, *et al*. Sharing of Individual Patient-Level Data by Trialists of Randomized Clinical Trials of Pharmacological Treatments for COVID-19. *JAMA* 2023; **329**: 1695-7. doi:10.1001/jama.2023.4590. pmid:37010865CrossRefPubMedGoogle Scholar
 112. Butte AJ. Trials and Tribulations-11 Reasons Why We Need to Promote Clinical Trials Data Sharing. *JAMA Netw Open* 2021; **4**: e2035043. doi:10.1001/jamanetworkopen.2020.35043. pmid:33507252CrossRefPubMedGoogle Scholar

Web appendix 1: Supplementary material (<https://pmc.ncbi.nlm.nih.gov/articles/instance/11231881/bin/mana078524.ww1.pdf>)

Web appendix 2: Acknowledgments (<https://pmc.ncbi.nlm.nih.gov/articles/instance/11231881/bin/mana078524.ww2.pdf>)

Web appendix 3: Combined CONSORT-Surrogate extension checklist

(<https://pmc.ncbi.nlm.nih.gov/articles/instance/11231881/bin/mana078524.ww3.pdf>)

Video 1: Surrogate outcomes