【実践・調査報告】

# エビデンスに基づく政策立案（EBPM）と研究論文の質改善のための報告ガイドライン

正木　朋也
国際協力機構／北里大学
masakit-tky@umin.ac.jp

津谷　喜一郎
東京有明医療大学／東京大学
tsutanik@gmail.com

## 要　約

日本では2010年代後半以降、公的セクターでの「エビデンスに基づく政策立案」（evidence-based policy making: EBPM）の方針が出された。医学領域では1996年にランダム化比較試験（randomized controlled trial: RCT）論文の質向上のためのCONSORT声明が開発された。報告ガイドラインの嚆矢となる。その第3版の全25項目からなるCONSORT 2010は16カ国語に翻訳され世界で広く使われている。また、観察研究（断面研究、後向き研究、前向き研究）やシステマティック・レビューなど基本的な研究デザインについて、さらにその拡張版としてクラスターランダム化や非薬物系などの種々の報告ガイドラインが開発され、その総数は400件を超す。これから研究計画を作成しようとする者にとり利用価値の高い、種々の報告ガイドラインをカバーするEQUATOR Networkが2008年に設立され広く用いられている。EBPMに報告ガイドラインを使う事はそれなりの実現可能性をもつ。

## キーワード

EBPM、報告ガイドライン、CONSORT、STROBE、EQUATOR Network

## 1. はじめに

私たちは本誌に過去2回「エビデンスに基づく医療」（evidence-based medicine: EBM）についてその展開を含めて報告してきた。ここでは2006年の論文（正木・津谷 2006）を第1報、2010年の論文（正木・津谷2010）を第2報と呼ぶこととする。それ以後の10年間でEBMの関連領域で起きたことが2つある。

1つは、日本の公的セクターでのEBPM（evidence-based policy making、エビデンスに基づく政策立案）の動きである。2017年5月、首相官邸が招集した「統計改革推進会議」の最終とりまとめ[1]が公表された。翌2018年10月には、「EBPM（エビデンスに基づく政策立案）に関する有識者との意見交換会報告（議論の整理と課題等）」[2]が公表され、全領域で公的なファンドを用いた研究が要請される時代となった。

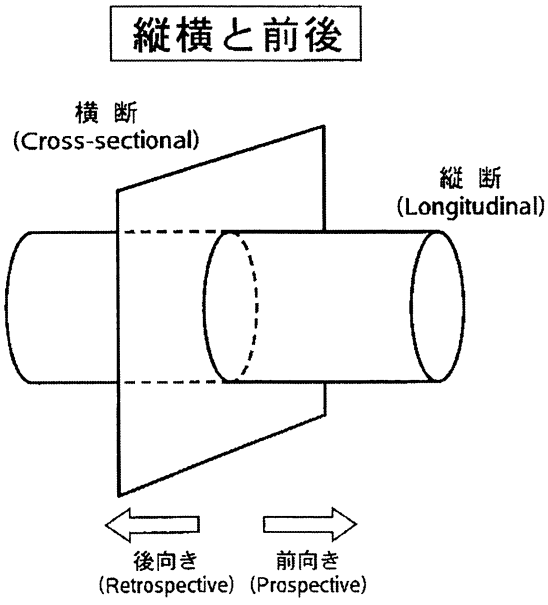
著者の一人、津谷は『いろいろな分野のエビデンス：温泉から国際援助までの多岐にわたるRCTやシステマティック・レビュー』を刊行した（津谷 2015）。その全12章のうち、国際援助と教育の章を除き、他の10章はすべて医学領域であった（第1報と第2報では「保健医療領域」の用語を用いた。本稿ではより分かりやすい「医学」を広義のものとして用いることとする[3]）。またそ

れらの多くは非政府系の活動であった。政府系、すなわち公的セクターでのEBPMは、研究者と納税者の双方の立場から質の高い研究が望まれる。

もう1つは、論文などの質を高めるための「報告ガイドライン」(reporting guidelines: RG) の多様な領域での数多くの発行とその使用である。ランダム化比較試験 (randomized controlled trial: RCT) に対する1996年のCONSORT (以降CONSORT声明をこう略す。他のRGについても同様) 第1版は、私たちの第1報で簡単に触れられた。また2001年の第2版は、第2報で若干紹介された。だが2010年3月に大きく改訂された第3版は、第2報が同時期の2010年3月発行であったことから紹介されていなかった。このCONSORT第3版は"CONSORT 2010"として知られ、その後多くの拡張版 (Extension) の作成が加速した。

一方、第2報で述べたように、1) コホート研究 (cohort study)、2) ケース・コントロール研究 (case-control study)、の時間軸をもつ2つの縦断研究 (longitudinal study) と、時間軸を持たない、3) 横断研究 (cross sectional study) の3つをカバーするSTROBE (von Elmら2007, 上岡 2008)、またその「解説と詳細」(explanation and elaboration: E&E, Vandenbrouckeら2007、日本語訳 福原ら2009) が2007年に発行された。この3つの研究デザインの関係を図1に示す (中嶋ら 2000)。

図1　時間軸を含めた研究デザインの整理




本稿では、第1に、日本でのEBPMの動向を念頭におき、まずバイアスの少ない、すなわちより正確 (accurate) な結果を得られるRCTのためのCONSORTについてその歴史的発展と拡張版などを含める。第2に、観察研究のSTROBEの3つの研究デザインのなかでケース・コントロール (case-control) 研究の部分のチェックリストを紹介する。第3に、多様な機能を持つEQUATOR NetworkとICMJEの動きについて述べ、第4に、医学の領域で開発されたこれらの手法やシステムのEBPMへの技術移転の実現性 (feasibility) について論ずる。

## 2. CONSORT

### (1) CONSORTの誕生

1996年のCONSORT (Beggら 1996;日本語訳: 津谷ら 1997) が注目されたのはその画期性とともに、周到にそのRGが有効かどうかの検証が行われたことが影響している。1996年にCONSORTが発表される前に公表されたRCT論文の質と、発表後に公表されたRCT論文の質を、世界的に著名な4つの総合医学雑誌を対象にして比較したのである。当時、すぐにCONSORTを論文投稿の条件とした3雑誌 (*BMJ, Lancet, JAMA*) は、以後、統計学的に有意に論文の質が向上した。一方でその時点ではCONSORTを投稿の条件にしていなかった*NEJM*では質の向上がなかったことを報告した (Moherら 2001)。

RCT論文の質評価のチェックリストの開発は1960年代から始まる。そのうち最も引用回数の多いのは予備的解析ではChalmers (1981) であった (津谷・佐久間 1993)。1992年に始まるコクラン共同計画 (The Cochrane Collaboration: CC) の初期[4]には、RCT論文の質評価に基づく選択に様々なチェックリストが使われていた。1996年の"Jadad Scale"は、ランダム化 (randomization)、ブラインディング (blinding)、除外・脱落 (withdrawal and dropouts) の3項目から成り立つ簡便なものであった (Jadadら 1996)。当時のコクランハンドブック (ver. 4.2.6) にも引用された。だがその後、単純すぎる、ブラインディングに重きがなされ

すぎている（Clarkら 1999）－などの理由で使われなくなる。

コクランレビューのアプリであるReview Manager（RevMan）は最初の版が1993年にリリースされた（Chandlerら 2013）。その2008年のversion 5（RevMan.5）に導入されたのが“Risk of Bias”toolのグラフである。そこでは7つの項目で各RCT論文の質を評価しその結果を棒グラフで表示している（Higginsら2011）。

### (2) CONSORT 2010の内容

RGは進化していくものである。1996年の第1版のチェックリストは項目にナンバリングがなかった。2001年の第2版は内容の改訂とともにナンバリングがなされ全22項目となった（Moherら 2001;津谷ら 2002）。2010年の現行の第3版ではチェックリストの項目数が25項目へと増加した（Schulzら 2010；日本語訳: 津谷ら 2010）。表2に示すこの3項目の追加には、臨床試験 を取り巻く環境の大きな変化が反映されている。

第23項目の「臨床試験登録」は、最初の試験対象者をリクルートする前に試験内容を項目毎にレジストリ（Registry）[5]に登録しそのレジストリの名称と登録番号を記入することである。Primary RegistryやData Providerは、日本を含めて現在世界17か国・地域に存在する。それらの試験情報はWHO-ICRPで統合されて2006年からすべて検索することができる。なお、WHOが必須とするdata setは当初の20項目から2017年に24項目に増加し2019年をゴールとするものである（津谷 2017a、2017b）。第24項目は、臨床試験に参加した個人のデータシェアリング・プラン（individual data sharing plan: IDSP）も含まれる（津谷 2018a、2018b）。

これの臨床試験登録の主たる目的は、結果が悪いと発表しない、または本来の1次アウトカム指標がよくない場合、2次アウトカム指標とされたものからよいものをあたかも1次アウトカム指標であるかのように報告する「選択的報告」（selective reporting）を含む「パブリケーションバイアス」を避けるためのものである（Macleodら 2014; 奥村ら 2014）。つまり結果がよくとも悪くともそれをプロトコール通りに解析し公開すべきだという考えである。それには、多くの患者の代表として試験に参加した患者から得られた結果は患者に戻されるべき公共財（public goods）であるという考え方がある[6]。

第24項目の「プロトコール入手可能性」は、プロトコールが遵守されず、上記のように本来とは異なった解析がなされる問題が顕在化したことに対応する。上記の選択バイアス防止の一手段である。

第25項目の「資金提供者」（funder）は、研究資金の提供主体が試験結果に影響する、つまり企業資金の場合、企業に有利な結果が報告されやすい、または不利な結果が報告されにくい、というエビデンスに基づく（Melander 2003）。なお、資金提供者とスポンサー（sponsor）は別のものであることに留意されたい（津谷・澤田 2007）。スポンサーの定義は3つの要素からなり分かりづらい[7]。

CONSORTのもう1つの重要な要素はフローチャートであり、図2に示す。このフローチャートは、先のJadad Scaleの第3項目の「除外・脱落」に関する事項を、より認知しやすい図形にしたものと理解できる。

なお、CONSORT声明は3つの版とも「解説と詳細」（explanation & elaboration）が発行され free access となっている。第2版のみ日本語訳が存在する（Altmanら 2001; 津谷・栗原 2002）

### (3) CONSORTの拡張版

表1に本稿執筆時点でのCONSORTのExtensionを、年代順に整理して示した。

ここで、SPIRIT（standard protocol items: recommendations for interventional trials）を紹介しよう（Chanら 2013、日本語訳 折笠ら 2017）。臨床試験が終わり、原稿を書く段になり、CONSORTのチェックリストに記載すべき情報が欠損していることが分かることがある。多くはプロトコールが不十分なためである。そこでプロトコール段階のチェックリストとして開発されたものである。大変有用なツールである。RCTの実施の前にprotocolを作成するのであるから「前向き」の「拡張版」ともいえ、後述するEQUATOR NetworkではCONSORTのカテゴリーに入っている。

表1　CONSORTのRCT領域に限ったExtensionのいろいろ

| | 試験デザイン | 介入内容 | データ |
|---|---|---|---|
| 2001 | ・クラスター(第1版)(日) | | |
| 2002 | | ・鍼(第1版)(日) | |
| 2004 | | | ・害(日) |
| 2006 | | ・ハーブ | |
| 2008 | ・実践的(日) | ・非薬物介入(第1版)(日) | ・抄録(日) |
| 2010 | | ・鍼(第2版)(日) | |
| 2012 | ・クラスター(第2版)<br>・非劣性・同等性 | | |
| 2013 | ・SPIRIT(プロトコール用) | | ・患者報告アウトカム(PRO)(日) |
| 2014 | | ・不妊治療 | |
| 2015 | ・N-of-1 | ・歯科矯正 | |
| 2016 | ・パイロット・実現可能性 | | |
| 2017 | ・個人内 | ・中薬複方<br>・非薬物介入(第2版)(日) | |
| 2018 | ・段階的割り込み型クラスター | ・社会的・心理的介入 | |
| 2019 | ・多群並行群間<br>・クロスオーバー<br>・N-of-1 TCM(中医学) | | |

(注)(日)日本語訳あり
(出所)津谷ら(2019)の整理にEQUATOR Networkをもとに更新(2019-10-12)

一部のRGでは、その理解を促すための「解説と詳細」も作成され多くはfree accessが可能である。これらRGやE&Eの一部は日本語にも翻訳されており、雑誌での公表や書籍として出版されている（中山・津谷 2008; 2016)。各国語訳されたRGやE&Eは、同じくEQUATOR Networkのwebsiteからpdfを入手できる。日本語訳についてはJPT Onlineが、それらをよくカバーしていると思われる。

### (4) CONSORTチェックリストとフローチャートの使い方

投稿者にとっての実際の使い方としては、論文を作成し投稿する者はCONSORTのチェックリストのフォーム（表2）に各項目の記載された原稿の「報告頁」を記載したものと、フローチャート（図2）にそれぞれの患者数を記載したものの2つを、投稿先へ原稿と共にemailなどで送る。これのチェックリストやフローチャートはpdfとWord版が各websiteにあり、これらをdownloadしてそこに記入すればよい。

それを編集委員会のスタッフが事務的にチェックし、ついで編集長や査読者（reviewer）にもそれらが原稿とともに送られて、各項目が実際に記載されているのかをチェックする。

フローチャートは論文そのものに掲載される。一方、チェックリストは多くはOpen access雑誌では、supplementary informationとして公開されることとなる。

例としてonline journalで、free accessとなっている*PLoS ONE*誌に掲載され、鍼の領域で2017年度の代田賞を獲得した、会陰部の夜間のかゆみに対するローラー鍼の二重盲検RCTの例をあげる(Iimuraら 2016)。ここから、チェックリストは、原稿の報告ページ番号を記載するが、電子出版の論文においてはページ番号が変わり、項目名で示されていることが分かる。またプロトコールや倫理委員会の審査記録などはこのsupplementary informationに収載され公開されている。

### (5) 非薬物介入とクラスターRCT

EBPMでの利用が期待されるCONSORT ExtensionのRGで社会科学の領域での使用が期待されるものが2つある。

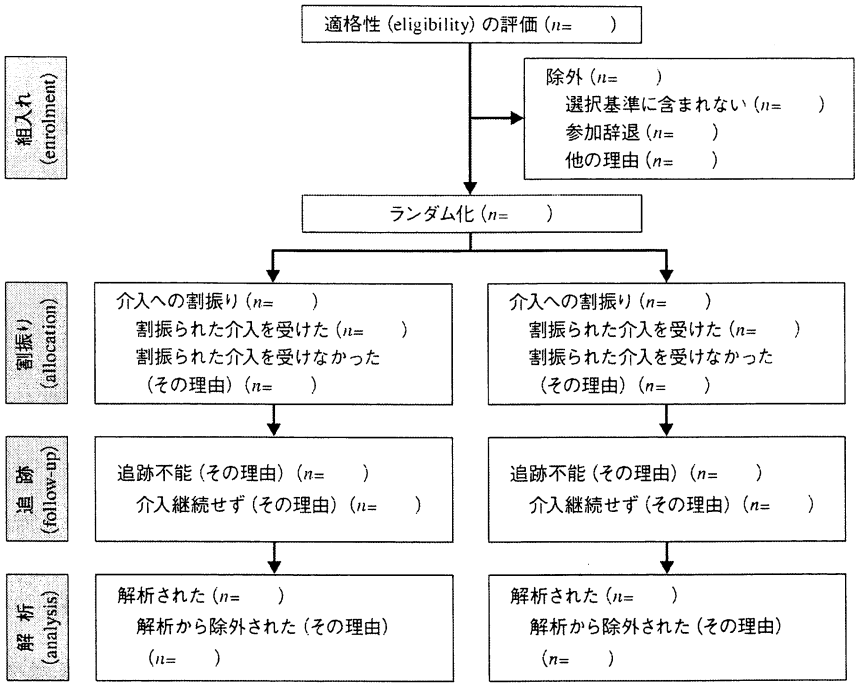
1つは、CONSORT非薬物介入（CONSORT NPT (non-pharmacologic treatment )）である（初版Boutronら 2008a、2008b、日本語訳 上岡・津谷 2016、第2版 Boutronら2017、日本語訳 津谷ら 2019)。非薬物介入の具体的なものとして、手術、手技（たとえば血管形成術などの技術的処置)、デバイス（ペースメーカーなど)、リハビリテーション、理学療法、行動療法、心理療法などを

## 表2　CONSORT 2010チェックリスト

| 章／トピック (Section/Topic) | 項目番号 (Item No) | チェックリスト項目 (Checklist Item) | 報告頁 (Reported on page No) |
|---|---|---|---|
| タイトル・抄録 (Title and Abstract) | | | |
| | 1a | タイトルにランダム化比較試験であることを記載。 | |
| | 1b | 試験デザイン(trial design)，方法(method)，結果(result)，結論(conclusion)の構造化抄録(詳細は「雑誌および会議録でのランダム化試験の抄録に対するCONSORT声明」[21, 31]を参照)。 | |
| はじめに (Introduction) | | | |
| 背景・目的 (Background and Objective) | 2a | 科学的背景と論拠(rationale)の説明。 | |
| | 2b | 特定の目的または仮説(hypothesis)。 | |
| 方法 (Method) | | | |
| 試験デザイン (Trial Design) | 3a | 試験デザインの記述(並行群間，要因分析など)，割付け比を含む。 | |
| | 3b | 試験開始後の方法上の重要な変更(適格基準 eligibility criteria など)とその理由。 | |
| 参加者 (Participant) | 4a | 参加者の適格基準(eligibility criteria)。 | |
| | 4b | データが収集されたセッティング(setting)と場所。 | |
| 介入 (Intervention) | 5 | 再現可能となるような詳細な各群の介入。実際にいつどのように実施されたかを含む。 | |
| アウトカム (Outcome) | 6a | 事前に特定され明確に定義された主要・副次的アウトカム評価項目。いつどのように評価されたかを含む。 | |
| | 6b | 試験開始後のアウトカムの変更とその理由。 | |
| 症例数 (Sample size) | 7a | どのように目標症例数が決められたか。 | |
| | 7b | あてはまる場合には，中間解析と中止基準の説明。 | |
| ランダム化 (Randomization) | | | |
| 順番の作成 (Sequence generation) | 8a | 割振り(allocation)順番を作成(generate)した方法。 | |
| | 8b | 割振りのタイプ：制限の詳細(ブロック化，ブロックサイズなど)。 | |
| 割振りの隠蔵機構 (Allocation concealment mechanism) | 9 | ランダム割振り順番の実施に用いられた機構(番号付き容器など)，各群の割付けが終了するまで割振り順番が隠蔵されていたかどうかの記述。 | |
| 実施 (Implementation) | 10 | 誰が割振り順番を作成したか，誰が参加者を組入れ(enrollment)たか，誰が参加者を各群に割付けた(assign)か。 | |
| ブラインディング (Blinding) | 11a | ブラインド化されていた場合，介入に割付け後，誰がどのようにブラインドかされていたか(参加者，介入実施者，アウトカムの評価者など)。 | |
| | 11b | 関連する場合，介入の類似性の記述。 | |
| 統計学的手法 (Statistical method) | 12a | 主要・副次的アウトカムの群間比較に用いられた統計学的手法。 | |
| | 12b | サブグループ解析や調整解析のような追加的解析の手法。 | |
| 結果 (Results) | | | |
| 参加者の流れ (Participant flow) (フローチャートを強く推奨) | 13a | 各群について，ランダム割付けされた人数，意図された治療を受けた人数，主要アウトカムの解析に用いられた人数の記述。 | |
| | 13b | 各群について，追跡不能例とランダム化後の除外例を理由とともに記述。 | |
| 募集 (Recruitment) | 14a | 参加者の募集期間と追跡期間を特定する日付。 | |
| | 14b | 試験が終了または中止した理由。 | |
| ベースライン・データ (Baseline data) | 15 | 各群のベースラインにおける人口統計学的(demographic)，臨床的な特性を示す表。 | |
| 解析された人数 (Number analyzed) | 16 | 各群について，各解析における参加者数(分母)，解析が元の割付け群によるものであるか。 | |
| アウトカムと推定 (Outcome and estimation) | 17a | 主要・副次的アウトカムのそれぞれについて，各群の結果，介入のエフェクト・サイズの推定とその精度(95%信頼区間など)。 | |
| | 17b | 2項アウトカムについては，絶対エフェクト・サイズと相対エフェクト・サイズの両方を記載することが推奨される。 | |
| 補助的解析 (Ancillary analysis) | 18 | サブグループ解析や調整解析を含む，実施した他の解析の結果。事前に特定された解析と探索的解析を区別する。 | |
| 害 (Harm) | 19 | 各群のすべての重要な害(harm)または意図しない効果(詳細は「ランダム化試験における害のよりよい報告：CONSORT声明の拡張」[28]を参照)。 | |
| 考察 (Discussion) | | | |
| 限界 (Limitation) | 20 | 試験の限界，可能性のあるバイアスや精度低下の原因，関連する場合は解析の多重性の原因を記載。 | |
| 一般化可能 (Generalisability) | 21 | 試験結果の一般化可能性(外的妥当性，適用性)。 | |
| 解釈 (Interpretation) | 22 | 結果の解釈，有益性と有害性のバランス，他の関連するエビデンス。 | |
| その他の情報 (Other information) | | | |
| 登録 (Registration) | 23 | 登録番号と試験登録名。 | |
| プロトコール (Protocol) | 24 | 可能であれば，完全なプロトコールの入手方法。 | |
| 資金提供者 (Funding) | 25 | 資金提供者と他の支援者(薬剤の供給者など)，資金提供者の役割。 | |

（注）＊標準2010年の改定で23、24、25の3項目が追加された。
（出所）津谷ら（2010）

図2　2群間並行ランダム化比較試験の各段階の過程を示すフローチャート
（組み入れ、介入への割振り、追跡、データ解析）




（出所）津谷ら（2010）

含み、幅広い領域をカバーする。

このCONSORT NPT Extensionは主に精神科医と外科医によって作られたものである。“Treatment”は「治療」だけではなく、「処置」や刑務所での「矯正」の訳がある。このRGは社会科学領域でもそのまま適用可能なRGであろう。EBPMの現場で行われるインパクト評価の計画と報告書作成にも有益である[8]。そこでは非薬物介入の複雑さ、施設とケア提供者の専門性の影響、ブラインド化の困難さも論じられている。

EBPMを推進しようとしている方は、RCTの論文原稿作成時ではなく、まずプロトコール作成時に先のSPIRITとこのCONSORT NPTを使用することが勧められる。

2つめは、クラスターRCTへの拡張版である（Campbellら 2004、2012）。ここには、同時に抄録の書き方のExtension（Hopewellら 2008）も含まれ、その全てのチェックリストが記載されている。10番目の実施（Implementation）の項目は10a、10bおよび10cに分割されている。

クラスターRCTはランダム割付け（random allocation またはrandomization）を、集団（cluster）を単位として行うものである。2009年頃より北米と西欧諸国を中心として世界的にクラスターRCTの利用が増えつつある（増島ら 2013）。なお、クラスターランダム化と他のランダム化の方法などの基本事項については他書を参照されたい（佐藤 2007）。

## 3. STROBEのケース・コントロール研究

「疫学研究における観察研究報告の質改善」（STROBE: strengthening of reporting of observational studies in epidemiological studies）のRGである。STROBEは、観察研究がその主体ともいえるEBPMでの応用が大いに期待される。観察研究は疫学研究から介入研究である臨床試験を除いたものである。

疫学（epidemiology）はその漢字表記から日本では医学領域であると誤解されがちである。だが、epiは周辺・辺縁の、-demosは人間・集団の、-logosは学問であり、人間を集団としてみる観点からあらゆる因果関係を明らかにしようとする

## 表3　STROBE: ケース・コントロール研究の報告において記載すべき項目のチェックリスト

| | No | 推奨 | 報告頁 |
|---|---|---|---|
| タイトル・抄録 [title and abstract] | 1 | (a) タイトルまたは抄録のなかで、試験デザインを一般に用いられる用語で明示する。<br>(b) 抄録では、研究で行われたことと明らかにされたことについて、十分な情報を含み、かつバランスのよい要約を記載する。 | |
| はじめに [introduction] | | | |
| 背景 [background]/論拠 [rationale] | 2 | 研究の科学的な背景と論拠を説明する。 | |
| 目的 [objective] | 3 | 特定の仮説を含む目的を明記する。 | |
| 方法 [methods] | | | |
| 研究デザイン [study design] | 4 | 研究デザインの重要な要素を論文のはじめの [early] 部分で示す。 | |
| セッティング [setting] | 5 | セッティング、実施場所のほか、基準となる日付については、登録、曝露 [exposure]、追跡、データ収集の期間を含めて明記する。 | |
| 参加者 [participant] | 6 | (a) 適格基準、参加者の母集団、ケース [case] の確定方法とコントロール [control] の選択方法を示す。ケースとコントロールの選択における論拠を示す。<br>(b) マッチング研究 [matched study] の場合、マッチングの基準、ケースあたりのコントロールの人数を記載する。 | |
| 変数 [variable] | 7 | すべてのアウトカム、曝露、予測因子 [predictor]、潜在的交絡因子 [potential confounder]、潜在的な効果修飾因子 [effect modifier] を明確に定義する。該当する場合は、診断方法を示す。 | |
| データ源 [data source]/測定方法 [measurement] | 8* | 関連する各因子に対して、データ源、測定・評価方法の詳細を示す。二つ以上の群がある場合は、測定方法の比較可能性 [comparability] を明記する。 | |
| バイアス [bias] | 9 | 潜在的なバイアス源に対応するためにとられた措置があればすべて示す。 | |
| 研究サイズ [study size] | 10 | 研究サイズ [訳者注：観察対象者数] がどのように算出されたかを説明する。 | |
| 量的変数 [quantitative variable] | 11 | 量的変数の分析方法を説明する。該当する場合は、どのグルーピング [grouping] がなぜ選ばれたかを記載する。 | |
| 統計・分析方法 [statistical method] | 12 | (a) 交絡因子の調整に用いた方法を含め、すべての統計学的方法を示す。<br>(b) サブグループと相互作用 [interaction] の検討に用いたすべての方法を示す。<br>(c) 欠損データ [missing data] をどのように扱ったかを説明する。<br>(d) 該当する場合は、ケースとコントロールのマッチングをどのように行ったかを説明する。<br>(e) あらゆる感度分析 [sensitivity analysis] の方法を示す。 | |
| 結果 [result] | | | |
| 参加者 [participant] | 13* | (a) 研究の各段階における人数を示す（例：潜在的な適格 [eligible] 者数、適格性が調査された数、適格と確認された数、研究に組入れられた数、フォローアップを完了した数、分析された数）。<br>(b) 各段階での非参加者の理由を示す。<br>(c) フローチャートによる記載を考慮する。 | |
| 記述的データ [descriptive data] | 14* | (a) 参加者の特徴（例：人口統計学的、臨床的、社会学的特徴）と曝露や潜在的交絡因子の情報を示す。<br>(b) それぞれの変数について、データが欠損した参加者数を記載する。 | |
| アウトカムデータ [outcome data] | 15* | 各曝露カテゴリー内の数、または曝露の集約尺度を示す。 | |
| おもな結果 [main result] | 16 | (a) 調整前 [unadjust] の推定値と、該当する場合は交絡因子での調整後の推定値、そしてそれらの精度（例：95%信頼区間）を記述する。どの交絡因子が、なぜ調整されたかを明確にする。<br>(b) 連続変数 [continuous variable] がカテゴリー化されているときは、カテゴリー境界 [category bordary] を報告する。<br>(c) 意味のある [relevant] 場合は、相対リスク [relative risk] を、意味をもつ期間の絶対リスク [absolute risk] に換算することを考慮する。 | |
| 他の解析 [other analysis] | 17 | その他に行われたすべての分析（例：サブグループと相互作用の解析や感度分析）の結果を報告する。 | |
| 考察 [discussion] | | | |
| 鍵となる結果 [key result] | 18 | 研究目的に関しての鍵となる結果を要約する。 | |
| 限界 [limitation] | 19 | 潜在的なバイアスや精度の問題を考慮して、研究の限界を議論する。潜在的バイアスの方向性と大きさを議論する。 | |
| 解釈 [interpretation] | 20 | 目的、限界、解析の多重性 [multiplicity]、同様の研究で得られた結果やその他の関連するエビデンスを考慮し、慎重で総合的な結果の解釈を記載する。 | |
| 一般化可能性 [generalisability] | 21 | 研究結果の一般化可能性（外的妥当性 [external validity]）を議論する。 | |
| その他の情報 [other information] | | | |
| 研究の財源 [funding] | 22 | 研究の資金源、本研究における資金提供者 [funder] の役割を示す。該当する場合には、現在の研究の元となる研究 [original study] についても同様に示す。 | |

*ケースとコントロールに分けて記述する。

（出所）上岡ら（2014、p.s116 表2）

学問であり、社会科学系の領域においても相補的に活用できる学問体系といえる。

表4に、日本で、すでに電子媒体などに蓄積されている過去のデータが使える「後向き」(retrospective) デザインであるケース・コントロール研究チェックリストを示した。なお、時間軸を含まない横断研究では原則的に因果関係の推論は不可能である。

## 4. EQUATOR NetworkとICMJE

EQUATOR Networkはこれまで述べてきたRGの動きを背景として2006年に設立されたネットワークである。以下の4つの目的を持つ。1) 良質な研究報告（論文）の重要性についての関心の向上、2) 各種研究デザインに対するよいRGの開発、流通（dissemination)、実施、の支援、3) 医学論文の質のモニタリング、4) 論文の質にインパクトを与える事項に関する－である (Simera 2009)。

臨床研究を計画し、また実施した後に投稿しようとする研究者、さらに編集者などのユーザーにとってはEQUATOR のwebsiteが大いに役立つ。2019年10月中旬時点でそのwebsiteは表4のように全領域を11項目に分類している。

そこには全419件のRGが存在する。対応する個々のRGにアクセス可能とし、そのExtension版など関連するガイドラインがリストされ、その多くはfree access の雑誌などにリンクされている。

このEQUATORと近い関係にあるのが国際医学雑誌編集者委員会 (international committee for medical journal editors: ICMJE) である。書誌情報の記載法の標準化を行った1978年のVancouver Groupをその母体とする。

1979年から発行されてきた "Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Medical Journal" (Uniform Requirementsと略されてきた、医学生物学雑誌の一般投稿規定）は数年に一度の改訂をして、世界の多くの医学雑誌の投稿規定のモデルとなり、医学雑誌の論文の質向上に貢献してきた。

ICMJE Recommendationsにfollow（従う）と宣言している雑誌は5,000誌以上ありリストが公開されている[9]。

これまで述べてきたCONSORTはその1997年の改訂で取り込まれた (Huthら 2004)。さらに利益相反に関わる記述の導入 (Fontanarosa 2010) などがなされてきた。また医学メディアを取り巻く環境の急速な変化を反映して2010年代となり改訂が多くなった[10]。

2013年8月の改訂でその名称は "Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals" (ICMJE

表4　試験の種類毎の報告ガイドライン

| 研究方法など | ガイドラインの略称 | Extension など |
|---|---|---|
| ランダム化試験 | CONSORT | 20 の拡張・関連版 (表 2 参照) |
| 観察研究 | STROBE | 11 の拡張・関連版 |
| システマティック・レビュー | PRISMA | 9 つの拡張・関連版 |
| 研究計画書 | SPIRIT | PRISMA-P |
| 診断／予後研究 | STARD | TRIPOD |
| 症例報告 | CARE | 3 つの拡張・関連版 |
| 臨床実践ガイドライン | AGREE | RIGHT |
| 質的研究 | SRQR | COREQ |
| 動物前臨床研究 | ARRIVE | |
| 質改善研究 | SQUIRE | |
| 経済評価 | CHEERS | |

（出所）EQUATOR Network. Webページを元に著者作成（閲覧 2019-10-12）

Recommendations）と変更された。"Requirements"から"Recommendations"とやや弱くなった形である。

RGは先に述べたように1997年とCONSORT発行の翌年に取り入れられたが、その後種々のRGが開発されたのに対応して、ICMJE Recommendationsの最新の2019年12月版にはこれまで述べてきたCONSORT、STROBE、さらにPRISMA（システマティック・レビュー用）、STARD（診断の正確性）が含まれるとともに、EQUATOR Networkが"resource"として記載されている。

ICMJEは、英米の有力雑誌5誌と世界各地の雑誌9誌の計14誌[11]と、米国国立医学図書館（National Library of Medicine: NLM）とWorld Association of Medical Editors（WAME）で構成される小さい組織である。アジア地域としては現在、韓国の旗艦雑誌（national flag journal）といえる*Journal of Korean Medical Science*（*JKMS*）が入っており、日本からは入っていない。このICMJEは規約がない組織であるが世界の医学雑誌をリードする存在でその影響力は大きい。

いわばEQUATOR NetworkがRGの開発機関とすると、ICMJEはそれらをそのRecommendationsに取り入れることで世界の雑誌に影響を与える執行機関で、表裏をなしているといえる。

## 5. 報告ガイドラインのEBPMへの技術移転

### (1) 形式がすぎて創造性を制限するのではないかとの批判

まず、RCTそれ自体の不理解と実施上の議論があることは押さえておく必要がある（Cartwright 2010; Deaton 2018）。そのうえで、しばしばRGのようなチェックリストは、無機的・機械的で味気なく、研究者の創造性を制限するものだと批判される。看護の領域では「いまだ定式化されていないものの、看護学の新しい知識の創出につながるような、画期的な研究が行われにくくなる一因となるのではないか」という考えも述べられている（萱間 2015）。著者らも最初にそれがRGの普及の阻害要因となるのでないかと考えた。

だが研究者は好きで研究を行うだけではない。特にRCTなどの臨床試験は人を対象にした人体実験である。そう考えると、個々の試験参加者に対する個別的倫理（individual ethics）の面からも、社会への集合的倫理（collective ethics）の面からも、この種のRGに沿った項目に沿って報告することは大切である（Ioannidis 2014）。また先に述べたICMJE Requirementsの項目として入り、RGに沿って記載されていな限り原稿が採択（accept）されないという制度的システムができたことも、医学領域において一定の成果を収めた理由になろう。

これはEBPMの領域においても同様である。試験参加者はそれなりの犠牲とリスクを伴うものであるとの認識が必要である。

### (2) ELSIの問題

ELSIはEthics, Law and Social Implicationの略である。ここでは2つを取り上げる。

1つは、先に述べたクラスターRCTである。集団（cluster）からのインフォームド・コンセント（informed consent: IC）をいかに受けるべきか、また、介入の有無により試験参加者集団の利益と不利益の均衡（equipoise）をいかにとるべきかの問題がある。クラスターRCTを行う場合の倫理面からの質向上のためのガイドラインとしてオタワ声明があり（Weijerら 2012）、その後の議論とともに更新の提案がなされている（van der Graafら 2015）。

2つは、利害関係者の調停と相互に利益相反の生じ兼ねない関係者間の問題である（Chalmers 2014）。医療における意思決定は、受益者である患者のために、公衆衛生上の政策と保険財務政策上の「折り合い」をつけ「落としどころ」をみつけて制度・政策として適用することがある。種々の利害関係者（stakeholder）、すなわち1）患者（とその家族や支援団体）、2）医療技術の開発者または製造販売企業、3）医療機関関係者（医師、病院経営者）、4）保健政策担当者（厚生労働省ほか保険財政・制度にかかわる関係各機関）、それら相互の利益や負担を全体として調停することが必要となる。

この種の問題はELSIとして研究対象になって

きた。医学領域での事例が、EBPMの領域での問題解決に役立つことはあろう。

### (3) ビッグデータやリアル・ワールド・データから得られる解とエビデンス

グローバルな巨大情報産業群（Google, Apple, Facebook, Amazon, Microsoft: GAFAM）などが開発・適用している人工知能（AI）を用いた即時提案型の情報技術に基づくビジネスモデルの応用も一考には値する。しかし、ここで扱われる「即時対応に用いる情報やAIが導く場当たり的な最適解」は、個々の場やケースにおいてのみ有効なものであるので、EBMに始まった「社会的に価値あるエビデンスを蓄積精査し公開しながら未来につなげようとする立場やアプローチ」とは本質的に異なるであろう。

なお、Oxfordの疫学統計学者Richard Petoが共著者に入りNEJMに掲載された“The magic of randomization versus the myth of real-world evidence”はこの領域に関心ある者には必読の論である（Collins 2020）。そこではまずデータの量が多くなろうとも観察研究の弱点そのものは変わらないことが述べられている。

### (4) システマティック・レビューから1次研究への揺り戻し

2009年に「システマティック・レビューとメタアナリシスの質改善」（Preferred Reporting Items for Systematic Reviews and Meta-Analyses: PRISMA）のRGが公表された。またシステマティック・レビューのプロトコール作成のためのPRISMA-Pが2016年に公表された。

ところが2019年の夏、ショッキングな報告がなされた（Niforatosら 2019）。それは、米国医師学会誌内科学（*JAMA Internal Medicine*）にInvitation letterとして掲載された2ページの論考である。その内容はSRとMAの数（SRMA）とRCTの数およびその比（SRMA/RCT比）を1995年から2017年までの22年間の経過を推移として追ったものであった。

医療分野毎の傾向をみたグラフでは、2017年時点で既にその比が1を超えている（つまりSRがRCTより多い）分野や診療領域が散見される状況があることが示された。これは同じRCTが幾度も別のSRに組み込まれて評価されている状況を示している。

その原因として、非ランダム化比較試験を含むSRやMAの増加、また中国の研究者らの投稿がPubMed内で検索し得るMAの3分の1あったことなどが挙げられている。SRの数の増加は筆者らの調査においてその動向を指摘している（唐ら 2014）[12]。また検索対象がPubMedだけであることや手法上の限界があり、どの程度のSRMA/RCT比が妥当であるかなどの議論はできないとしている。

先のビッグデータやRWDへの懸念とともに、どうやら1990年代に始まるシステマティック・レビューの成長は少なくとも医学領域ではその量と質から陰りを見せ、1次研究、特に質の高いRCTが再び重要性を取り戻したようだ。EBPMにおいても常にエビデンスの質に留意する必要があろう。

### (5)「産官学」での意思決定のための文書作成の異同

これまで述べてきたようにRGの方法は「産官学」の「学」（academia）、特に医学の領域で発展してきた。他の領域へのその技術移転の実現可能性（feasibility）はどうだろうか？「官」の公的機関によるEBPMではそのプロセスの中で、雑誌に掲載される際の査読（review）がとられることはあるだろうか？なかには外部の査読付きの雑誌に投稿することもあろう。本誌は査読付きの雑誌である。

だが、行政機構の中では「査読」の用語はあまり一般的ではないかもしれない。しかし政策実施のための資料はその性格から、決して一人で作成されるものではなく執筆者の所属する組織内、また関係する他の組織の、複数の目を経て加筆修正されて、最終版は作成される。そこでの質管理にはRGが同様に利用される可能性が多いであろう。

医学以外の「学」の分野ではどうだろうか？社会科学の領域ではインパクト評価の実施数が増えつつある（Cameronら 2015）。また社会科学の重要な部門の1つである経済学においてすでに

RGが取り上げられている。

「公」ではどうであろうか？世界を見れば、2017年に世界銀行は、経済学研究は殆ど、研究デザインにおける重要な項目の詳細を報告していないために、選択（selection）、損耗（attrition）、実行（performance）、検出（detection）、報告（reporting）、サンプルサイズ（sample size）に関わる6種類のバイアスが生じている可能性があるとしている。そしてその対策としてCONSORTを利用することを推奨している（Ebleら 2017）。

政府間機関（intergovernmental organization）である国連や国連システムの1つである世界銀行はいわば公的機関である。日本の都道府県の議会での発言中の「エビデンス」の頻度はかなりばらつきがある[13]。一方でいくつかの地方自治体はすでにEBPMのための政策文書を作成している[14]。

## 6. おわりに

バイアスを小さくしエビデンスの強い方法を開発して適応するという枠組みは、社会改善（social betterment）の本質である。医学領域と他の領域に違いはない。EBMの発展の中で2010年代になり急速に進化したRGのうち2つを紹介した。またその開発・使用を支援するEQUATOR-Networkを紹介した。RGのEBPMへの技術移転の実現可能性はそれなりに存在することを論じた。そこでRGは単なる報告書作成時のマニュアルとしてではなく、その本質を理解し「賢く」計画段階からチェックリストも念頭にデザインを考え、実施可能なところは積極的に適用してゆくことが重要である。

### 謝辞

本稿作成に協力いただいた聖路加国際大学看護学研究科八重ゆかり氏に謝意を表する。

### 注記

1 「統計改革推進会議 最終取りまとめ」2017.5

2 総務省EBPMに関する有識者との意見交換会事務局、「EBPM（エビデンスに基づく政策立案）に関する有識者との意見交換会報告（議論の整理と課題等）」、2018.10

3 日本医学会は、基礎医学、社会医学、臨床医学の3つの部門をもつ。本稿では、歯学、薬学、看護、また多様なco-medicalの領域を含むものとして、他領域の関係者にも分かりやすい「医学」を用いる。

4 そのwebsiteには2019.10時点で8,107件のコクランレビューのサマリーがfree accessとなっている（2015.1.31より単にCochrane（コクラン）に名称変更しロゴマークも変わった。本稿ではこれまでどおり「コクラン共同計画」の日本語訳を用いた）。

5 臨床試験の登録システムのこと。国際的に合意された項目のWHO data setは現在14項目（ver.1.3.1）. UMIN-CTRやjRCTなど日本のものも含めて各国のRegistryの登録項目はそれよりも一般に多い。なお本稿では以降CONSORTの項目番号とWHO ICTROの項目番号の2種が記載され混乱しやすいが、一度、現物をご覧になると容易に区別できるであろう。

6 津谷喜一郎、臨床試験登録の意義と役割について、UMIN臨床試験登録システム シンポジウム（2005.2.2）

7 WHO GCP（1995）とICH GCP（1996）の定義の原文は、"The individual, organization, group or other legal entity which takes responsibility for initiating, managing and/or financing a study" である。ここで、3つの要素を "and/or" で関係づける表現は、そのまま訳すと日本語としてわかりにくい。そこで、以下がよいと思われる。「ある試験の企画、管理、資金のすべて、もしくはいずれかに責任を持つ、個人、組織、グループ、または他の法的主体」津谷喜一郎 編集ガイドラインとCOI. in: 第7回日本医学雑誌編集者会議（JAMJE）第7回シンポジウム「不正論文と編集者の責務」、slide 24

8 国際開発分野のインパクト評価のエビデンス情報の集積と提供を行うInternational Initiative for Impact Evaluation（3ie）は2009年より国際開発分野としてC2に合流した。

9 だが、雑誌リストの前書きにこの長いリストは完全なwhite listではなく、悪徳雑誌（predatory journal）も入っている可能性があるとしており、日本医学会も同様の注意喚起を行っている。

10 ICMJE Recommendationsの改訂履歴（Archives）はWeb上で確認可能である。

11 14の雑誌と2つの機関がある。5大雑誌は、通常、米国のNew England Journal of Medicine、JAMA（Journal of American Medical Associationの略だがJAMAが正式名称）、Annal of Internal Medicine －の3つと、イギリスのLancetとBMJ（こちらもBritish Medical Journalの略でBMJが正式名称）－の2つである。いずれにしろ、EBPMのために日本の各領域における指導的雑誌の質向上が期待され、その投稿規定にCONSORTなどの基本的RGの使用が含まれることが望まれる。

12 医学領域のPubMedにindexされる狭義のシステマティック・レビューの総数は2013年末で約75,000件と推計された（唐ら 2014）。世界の多くの研究方法論者（methodologist）が待っていたsystematic reviewのPublication Type（PT）が昨2019年にようやく作成された。PubMedでそのPTを用いてsystematic reviewを1950.1.1から2019.12.31までとして検索すると約124,000件、2019年の1年間で13,000件である。急速に大量のsystematic reviewが「つくら」れていることは間違いない。

13 2019年1月末時点の各都道府県の議会での「エビデンス」の用語の頻度分析「地方議会における EBPM（証拠に基づく政策立案）の現状と課題（堀江 2019）」の報告がある。

14「エビデンスに基づく政策展開の推進」調査研究チーム 最終報告書（平成31年3月）」（北海道調査研究チーム 2019）。

## 参考文献


家子直幸・小林庸平・松岡夏子・西尾真治（2016）「エビデンスで変わる政策形成～イギリスにおける「エビデンスに基づく政策」の動向、ランダム化比較試験による実証、及び日本への示唆～」、『三菱UFJリサーチ&コンサルティング政策研究レポート』、https://www.murc.jp/wp-content/uploads/2016/02/seiken_160212.pdf（閲覧日 2019-9-30）

奥村泰之・吉田和樹・清水沙友里（2014）「臨床疫学研究における報告の質向上に向けて：研究者の倫理.」、『精神科』、24（5）：551-557

上岡洋晴・津谷喜一郎（訳）（2008）「疫学における観察研究の強化（SROBE声明）：観察研究の報告に関するガイドライン」、中山健夫・津谷喜一郎編著（2016）『臨床研究と疫学研究のための国際ルール集Part2』、ライフサイエンス出版、202-209

上岡洋晴・中山健夫・津谷喜一郎（2014）「EQUATOR Networkから得られる，質の高い研究報告のための国際ルール－③ STROBE: 疫学における観察研究報告のためのルール」、『薬理と治療』、42（suppl.2）：s109-117

上岡洋晴・津谷喜一郎（訳）（2016）「CONSORTグループの方法とプロセス: 非薬物療法評価試験のための拡張版」、中山健夫・津谷喜一郎編著『臨床研究と疫学研究のための国際ルール集Part 2』、ライフサイエンス出版、70-9

萱間真美（2015）「論文の査読ガイドラインの構築に向けて－誰のためのガイドラインか」、『看護研究』、48（7）：642-7

佐藤俊哉（2007）「人間栄養学講座 連載 ランダム化臨床試験をするまえに【第4回】ランダム化の方法」、『栄養学雑誌』、65（5）：255-60

総務省（2018.10）『EBPM（エビデンスに基づく政策立案）に関する有識者との意見交換会報告（議論の整理と課題等）』、EBPMに関する有識者との意見交換会事務局 http://www.soumu.go.jp/main_content/000579366.pdf（閲覧日 2019-9-30）

統計改革推進会議（2017）『統計改革推進会議：最終とりまとめ』（平成29年5月19日統計改革推進会議決定）、https://www.kantei.go.jp/jp/singi/toukeikaikaku/pdf/saishu_honbun.pdf（閲覧日 2019-9-30）

中嶋宏（監修）・津谷喜一郎・坂巻弘之・山崎茂明（2000）『EBMのための情報戦略－エビデンスをつくる、つたえる、つかう』、中外医学社

中山健夫・津谷喜一郎（2008）『臨床研究と疫学研究のための国際ルール集』、ライフサイエンス出版

中山健夫・津谷喜一郎（2016）『臨床研究と疫学研究のための国際ルール集〈part2〉』、ライフサイエンス出版

津谷喜一郎・佐久間昭（1992）「臨床試験論文チェックリストのチェック－情報医学的観点から－」、（1993）『臨床薬理』、24（1）：49-50

津谷喜一郎・澤田弘（2007）「臨床試験のsponsorはスポンサーではない」、『臨床薬理』、38 Supple：S268

津谷喜一郎・元雄良治・中山健夫（2010）「CONSORT 2010声明：ランダム化並行群間比較試験報告のための最新版ガイドライン」、『薬理と治療』、38（11）：

939-947

津谷喜一郎（2010）「CONSORT 2010 訳者解説」、『薬理と治療』、38（11）：948-9

津谷喜一郎（2015）『いろいろな分野のエビデンス－温泉から国際援助までの多岐にわたるRCTやシステマティックレビュー』、ライフサイエンス出版

津谷喜一郎・中澤幾子・佐藤元（2017a）「WHO 臨床試験登録データセット（Version 1.2.1, Version 1.3）」、『薬理と治療』、45(12)：1905-10

津谷喜一郎（2017b）「WHO 臨床試験登録データセット解説」、『薬理と治療』、45(12)：1911-6

津谷喜一郎・中澤幾子・佐藤元.（2018a）「WHO 臨床試験登録データセット（Version 1.3）の背景となる4つの資料の日本語訳）」、『薬理と治療』、46(5)：669-682

津谷喜一郎・中澤幾子・佐藤元（2018b）「WHO 臨床試験登録データセット（Version 1.3）の背景となる4つの資料の解説」、『薬理と治療』、46(6)：965-712

津谷喜一郎・上岡洋晴・折笠秀樹・佐藤元（2019）「2017 CONSORT 非薬物介入版の紹介と解説－2017 CONSORT NPT Extension－」、『薬理と治療』、47(6)：865-84

津富宏訳・津谷喜一郎監訳（2008）「CONSORT声明: クラスターランダム化試験への拡張」、中山健夫・津谷喜一郎 編著（2008）『臨床研究と疫学研究のための国際ルール集』、ライフサイエンス出版、105-17

唐文涛・小島原典子・河合富士美・津谷喜一郎（2014）「診療ガイドラインとシステマティック・レビュー」、『薬理と治療』、42(3)：189-197

福原俊一・山口拓洋・山崎新・林野泰明・竹上未紗（2009）『観察的疫学研究報告の質改善（STROBE）のための声明：解説と詳細』、（2009年5月28日　日本語翻訳検討グループ翻訳）

正木朋也・津谷喜一郎（2006）「エビデンスに基づく医療（EBM）の系譜と方向性: 保健医療評価に果たすコクラン共同計画の役割と未来」、『日本評価研究』、6(1)：3-20

正木朋也・津谷喜一郎（2010）「エビデンスに基づく医療（EBM）の展開から学ぶもの－EBM普及に伴い経験した課題と解決策－」、『日本評価研究』、10（1）：3-16

増島祥・正木朋也・唐文涛・津谷喜一郎（2013）「日本と世界におけるクラスターランダム化比較試験の現状と課題」、『第34回日本臨床薬理学会学術総会』、東京

Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Group, f. t. C. (2001). The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration. *Annals of Internal Medicine*, 134(8), 663-694. [津谷喜一郎、栗原千絵子（訳）（2002）ランダム化比較試験報告のためのCONSORT声明改訂版－解説と詳細、『臨床評価』、30(1): 163-208]

Begg, C., Cho M., Eastwood, S., et al (1996). Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 276(8): 637-9. [日本語訳：津谷喜一郎・小島千枝訳（1997）「無作為化比較試験の報告の質を改善する方法CONSORT声明」、『*JAMA*〈日本語版〉1997年7月号』、74-7]

Boutron I., Moher D., Altman D.G., et.al, CONSORT Group. (2008a). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med*, 148: 295-309.

Boutron I., Moher D., Altman D.G., et.al, CONSORT Group. (2008b). Methods and processes of the CONSORT Group: example of an extension for trials assessing nonpharmacologic treatments. *Ann Intern Med*, 148:W60.

Boutron I.B., Altman D.G., Moher D., et.al for the CONSORT NPT Group. (2017). CONSORT Statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT Extension for nonpharmacological trial abstracts. *Ann Intern Med*, 167: 40.

Cameron, D.B., Mishra, A., Brown, A.N. (2015). The growth of impact evaluation for international development: how much have we learned? *Journal of Development Effectiveness* 8: 1-21.

Campbell M.K., Elbourne D.R., Altman D.G. (2004). CONSORT statement: extension to cluster randomised trials. *BMJ*; 328:702-8.

Campbell, M.K., Piaggio, G., Elbourne, D.R., et.al & Group, C. (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ*, 345, e5661. doi:10.1136/bmj. e5661

Cartwright, N., & Munro, E. (2010). The limitations of

randomized controlled trials in predicting effectiveness. *Journal of Evaluation in Clinical Practice*, 16(2), 260-266.

Chalmers, I., Bracken, M. B., Djulbegovic, B., et.al. (2014). How to increase value and reduce waste when research priorities are set. *Lancet*, 383 (9912), 156-165.

Chan, A.-W., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gøtzsche, P. C., Krleža-Jerić, K., Moher, D. (2013) . SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Annals of Internal Medicine*, 158(3), 200. [日本語訳: 折笠秀樹・津谷喜一郎・上岡洋晴 (2017)「SPIRIT 2013 声明：臨床試験のための標準的なプロトコール項目の規定」、『薬理と治療』245(12): 1895-1904]

Chan, A. W., Song, F., Vickers, A., et.al. (2014). Increasing value and reducing waste: addressing inaccessible research. *Lancet*, 383 (9913), 257-266. doi:10.1016/S0140-6736(13)62296-5

Chandler, J., Hopewell, S. (2013). Cochrane methods - twenty years experience in developing systematic review methods. *Systematic Review*, 2(1), 76.

Clark, H. D., Wells G. A., Huët C., et al. (1999). Assessing the quality of randomized trials: reliability of the Jadad scale. *Controlled Clinical Trials*, 20(5): 448-52.

Collins R, Browman L, Landray M, et. al(2020). The magic of randomization versus the myth of real-world evidence. *New England Journal of Medicine*. 382 : 674-678.

Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2-21.

Eble A., Boone P., Elbourne D. (2017). On minimizing the risk of bias in randomized controlled trials in economics. *The World Bank Economic Review*, 31(3): 687-707.

Fontanarosa, P. B., Flanagin, A., & DeAngelis, C. D. (2010). Implementation of the ICMJE form for reporting potential conflicts of interest. *JAMA*, 304(13), 1496.

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., et. al. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343: d5928.

Huth J., E., Case, K. (2004). The URM: Twenty-five years old. *Science Editor*, 27(1), 17-21.

Iimura K., Watanabe N., Masunaga K., et al. (2016). Effects of a gentle, self-administered stimulation of perineal skin for nocturia in elderly women: A randomized, placebo-controlled, double-blind crossover trial. *PLOS ONE*, 11(3), e0151726.

International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals (ICMJE Recommendations, formerly the Uniform Requirements for Manuscripts). http://www.icmje.org/recommendations/archives/ (閲覧日 2019-9-30)

Ioannidis, J. P., Greenland, S., Hlatky, M. A., et. al. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *Lancet*, 383(9912), 166-175.

Jadad A. R., Moore R. A., Carroll D., et .al. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary?. *Control Clin Trials*. 17(1), 1-12.

Macleod, M. R., Michie, S., Roberts, I., et.al. (2014). Biomedical research: increasing value, reducing waste. *Lancet*, 383(9912), 101-104.

Melander H., Ahlqvist-Rastad J., Meije G., et.al. (2003) . Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ*, 326, 1171-5.

Moher D., Schulz K. F., Altman D. for the CONSORT Group. (2001). The CONSORT statement: revised recommendations for improving the quality of parallel-group randomized trials. *JAMA*, 285(15), 1987-1991 [津谷喜一郎・小島千枝・中山健夫訳 (2002)「CONSORT 声明：ランダム化並行群間比較試験報告の質向上のための改定版勧告」、『JAMA〈日本語版〉2002年6月号』、74－7]

Moher D., Jones A., Lepage I., for the CONSORT Group. (2001). Use of the CONSORT statement and quality reports of randomized trials: a comparative before-and-after evaluation. *JAMA*, 285, 1992-5.

Moher, D., Hopewell, S., Schulz, K. F., et.al. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, c869.

Niforatos J. D., Weaver M., Johansen M. E. (Published online July 29, 2019). Assessment of publication trends of systematic reviews and randomized clinical trials, 1995 to 2017. *JAMA Intern Med*.

Schulz, K. F., Altman, D. G., Moher, D., & Group, C. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, c332. doi:10.1136/bmj. c332［日本語訳: 津谷喜一郎・元雄良治・中山健夫訳 (2010)「CONSORT声明: ランダム化並行群間比較試験報告のための最新版ガイドライン」、『薬理と治療』、38(11): 939-947］

Simera, I., Altman, D. G. (2009). Writing a research article that is "fit for purpose" : EQUATOR Network and reporting guidelines. *Evidence-based medicine*, 14(5), 132-4.

van der Graaf, R., Koffijberg, H., Grobbee, D. E., et al. (2015). The ethics of cluster-randomized trials requires further evaluation: a refinement of the Ottawa Statement. *J Clin Epidemiol*, 68(9), 1108-1114.

Vandenbroucke J. P, von Elm E., Altman D. G., et. al. STROBE Initiative (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *Epidemiology*,18, 805-835.

von Elm E., Altman D. G., Egger Mattjoas, et. al. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE). *PLoS Medicine*, 4(10), 1623-1627.

Weijer C., Grimshaw J.M., Eccles M.P., et al. (2012). The Ottawa statement on the ethical design and conduct of cluster randomized trials. *PLoS Med*, 9, e1001346.

(2020.5.25 受理)