

Series

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 —— ⑥

STARD：診断の精度に関する研究報告のためのルール

酒井 恭子* 中山 健夫**

EBM と論文執筆ガイドラインの充実

根拠に基づく医療 (evidence-based medicine: EBM) は、“最善の根拠，臨床的熟練，患者の価値観，患者個々の臨床状態と置かれている環境を統合すること”¹⁾とされ，現在では医療の意思決定の基本理念となっている。最善の根拠，すなわち best research evidence には，質の高い研究報告が不可欠であり，本シリーズでは EQUATOR (Enhancing the QUALity and Transparency Of health Research) Network に含まれる代表的な執筆ガイドラインである CONSORT (ランダム化比較試験)，STROBE (観察研究)，PRISMA (系統的レビュー) などを紹介してきた。今回は診断精度の研究報告における論文執筆のガイドラインとして，2003 年に発表された STARD (Standards for Reporting of Diagnostic Accuracy) 声明と，その改訂版である STARD 声明 2015 について解説する。

診断研究と執筆ガイドラインの必要性

診断検査の領域は変化が激しい。新しい検査が開発され，診断技術は次々と進化していく。現代の医療において，診断検査は診断や治療方針の意思決定の重要な鍵である。診断検査が，妥当性の低い評価のまま実地医療に導入されると，誤った診断・病態

把握という診療の正しい意思決定を根本から揺るがすことが起こりかねない。

一方で，診断精度に関連する研究は，デザインや実施についての批判的な情報の欠如が指摘されてきた^{2,3)}。コクラン共同計画の診断およびスクリーニング検査ワーキンググループは，検査に関する研究の方法論的な質の低さについて議論を積み重ね，介入研究の CONSORT 声明のように，診断精度の研究報告に含むべき項目のチェックリスト作成を提言した。

STARD 声明の誕生へ

1 STARD の作成

初版の STARD 声明は，科学者と学術編集者らのグループが診断精度に関する研究報告の質を向上させるため作成し，2003 年に発表した⁴⁾。研究報告に必要な内容が適切かつ十分記載されていれば，読者がその研究におけるバイアスに気づきやすくなり，その結果，的確に研究の内的妥当性を評価できるようになる。加えて結果の外的妥当性（一般化可能性 [generalisability]）についても，読者が自分自身のセッティングで，その報告の知見をどのように活用できるか，活用し慎重であるべきか，考察するのに役立つ⁵⁾。

*大阪府済生会吹田病院臨床検査科

**京都大学大学院医学研究科社会健康医学系専攻健康情報学

II STARD チェックリスト

STARD 声明 2003 には、著者が論文の執筆に際して、すべての関連情報が報告に含まれているかどうかを確認する 25 項目のチェックリストとフローチャートより成る。その後、STARD 声明 2015⁶⁾では、チェックリストは 30 項目に改訂された(表 1)。

1 診断研究の特徴

・リサーチクエスション、デザイン

診断確度に関する研究の基本構造は、対象状態を検出・予測するために、ある検査(1 つ以上の場合もある)を評価するかたちである。評価の対象はまず、その検査で対象状態を判別可能か否か、ということであり、リサーチクエスションは「ある検査結果が得られた患者は、そうでない検査結果を得た患者よりも対象状態にある可能性が高いか?」となる。次の段階は、検査 [test under evaluation] と標準検査 [reference standard] の双方から得たアウトカムの比較を行う。

・ばらつき・バイアス発生の可能性

検査確度の測定値は研究ごとに異なる。ばらつき [variability] は患者群、セッティング、対象状態の定義、検査プロトコルや陽性に関する判定基準などの違いを反映している。

リサーチクエスションとは異なる状況で検査を実施した場合、たとえば、病状が進行している患者で、病初期に行うスクリーニング検査を評価したり、診療所で用いる検査・機器を大学病院の専門科で評価したりすると、結果の妥当性を損なうバイアスが生じ得る。

2 STARD 声明の特徴

STARD 声明のチェックリストから、診断確度研究に独自の点を以下に挙げる。

・項目 1 タイトル・抄録

他の声明では、研究デザインを述べる (CONSORT 声明ではランダム化比較試験であることをタイトルで明示することを求めている) が、STARD 声明では研究の主題が検査確度であることを示すために、感度・特異度や薬物血中濃度-時間曲線下面積 (area under the blood concentration time curve: AUC) といった、検査研究独自の測定指標を含めるように指示している。

・項目 11 方法

その標準検査を選択した論拠を記載。

・項目 12a, 12b 方法

カットオフ値を設定した論拠を記載。

・項目 17 解析

対象予備軍から対象者を識別する場合の、診断確度のばらつきについて検討した結果を (その検討が探索的なものか、事前に特定したものであったかを区別して) 記載。

・項目 21a 結果 (参加者)

対象状態にある患者の疾病の重症度の分布を報告 (判定基準を定義する)。重症度の分布は感度・特異度に影響するので重要。

・項目 21b 結果 (参加者)

対象状態にない患者については、どのような診断を受けた患者かを報告。

・項目 22 結果 (検査結果)

目的検査を実施してから標準検査を行うまでの時間とその間に行われた治療を報告。

・項目 23 結果 (推定)

標準検査の結果別に目的検査の結果を表示。

III STARD 声明 2015 の改訂点

2015 年の改訂では、チェックリスト項目が再構成され (表 2)、フローチャートが変更された (図 1)。下記に 2015 年版で追加された項目を挙げる。これらの多くは、検査研究独自のものというより、妥当性の高い研究と報告を推奨し、研究不正を防ぐため、近年要求されている研究報告の在り方を反映したものである。

・抄録 Abstract

項目 2 構造化抄録のかたちで記述: 抄録が、研究デザイン・結果の key elements の検索に使用されることが多い。

・はじめに Introduction

項目 3 検査の目的と診療上の役割: その研究で対象とされた集団を明示することで、読者は自分自身のセッティングで、その報告の知見をどのよう利用できるか推測できる。

項目 4 研究の仮説: 研究の仮説が明確でないと、検査結果がさまざまに解釈されてしまい、その結果、結論の方向が定まらなくなってしまう (spin)。

表 1 STARD 声明 2015 チェックリスト

章・トピック [Section & Topic]	番号 No	項目 Item
タイトルあるいは抄録 [TITLE OR ABSTRACT]	1	診断確度の少なくとも1つの測定指標（感度，特異度，陽性的中率，曲線下面積など）を用いて，論文が診断確度の研究であると特定
抄 録 [ABSTRACT]	2	研究デザイン，方法，結果，結語を含む構造化された要旨（詳細は，STARD for Abstracts 参照）
はじめに [INTRODUCTION]	3	目的検査 [index test] の使用目的と診療上の役割を含む科学的，臨床的背景
	4	研究目的と仮説
方 法 [METHODS]		
研究デザイン [Study design]	5	データ収集が，目的検査および標準検査 [reference standard] を実施する前に行われたか（前向き [prospective] 研究）あるいは実施後か（後ろ向き [retrospective] 研究）を記述
参加者 [Participants]	6	選択基準 [eligibility criteria]
	7	症状，先の検査結果，あるいは登録組み入れ基準など，どのように参加者を選択したかを記述
	8	セッティング [setting] およびデータを収集した場所，時
	9	参加者は連続した [consecutive]，ランダム化された，あるいは適合する患者かどうかを記述
検査方法 [Test methods]	10a	目的検査において，再現性の方法記述が十分かを記述
	10b	標準検査において，再現性の方法記述が十分かを記述
	11	（もし代替検査があれば）その標準検査を選んだ論拠 [rationale]
	12a	目的検査の結果の単位，カットオフ値および/または分類に関する定義と根拠を記述
	12b	標準検査の結果の単位，カットオフ値および/または分類に関する定義と根拠を記述
	13a	目的検査を行う人また読者に，標準検査の結果と他の臨床情報が提供されているかを記述
	13b	標準検査の判読者に，目的検査の結果と他の臨床情報が提供されているかを記述
解 析 [Analysis]	14	診断確度の計算または比較に用いた方法
	15	目的検査あるいは標準検査で中間的 [intermediate] な結果が得られた時に，それをどう扱ったかを記述
	16	目的検査および標準検査の欠測値をどう扱ったかを記述
	17	対象予備軍から対象者を識別する場合の，診断確度のばらつきについての解析
	18	目的とされたサンプルサイズとその見積り方法
結 果 [RESULTS]		
参加者 [Participants]	19	参加者のフローチャートを報告
	20	参加者の人口統計的 [demographic] および臨床的特性の基準を報告
	21a	対象状態にある患者の疾病の重症度の分布を報告（判定基準を定義）
	21b	対象状態にない患者については他の診断を報告
	22	目的検査を実施してから標準検査を行うまでの時間，およびその間に行った治療を報告
検査結果 [Test results]	23	標準検査の結果別に目的検査の結果（あるいは分布）を表で示す
	24	診断確度の推定値と，統計的不確実性（例：95%信頼区間）を記述
	25	目的検査もしくは標準検査により生じた有害事象 [adverse event] を記述
考 察 [DISCUSSION]	26	バイアスの存在・統計学的不確実さ・一般化の可能性を含む研究の限界
	27	目的検査が対象とする集団と目的が考慮された，診療および研究に対する示唆
その他の情報 [OTHER INFORMATION]	28	研究登録の詳細（登録番号および氏名）
	29	正式な研究プロトコルのアクセス可能なところ
	30	資金源，その他のサポーターとその役割

表 2 STARD 声明 2015 で追加された項目の概要

番号 No	項目 Item	理由 Rationale
2	構造化抄録	抄録が、研究デザイン・結果の key elements の検索に使用されることが増えてきている
3	検査の使用目的と診療上の役割	研究が対象とした集団を明示することで、読者は自分たちのセッティングでの測定値の精度を推測できる
4	研究の仮説	定まった研究の仮説が定まっていなくて、検査結果は漫然とした対象に対して用いられ、その結果結論の方向が定まらなくなってしまう
18	サンプルサイズ	読者は、研究で推定された診断確度と診断力、および著者が目的とした参加者数を集めることに成功したかどうかを評価することを望んでいる
26・27	構造化された考察	望ましくない結論を導くことをさけるため、著者は、研究の限界（バイアスの存在、統計学的不確実さ、一般化の可能性）や、検査が対象とする集団と目的を考慮した示唆を書く
28	研究登録の詳細	新しい検査の確度の研究は臨床試験であり、研究開始前に臨床試験登録が必要である。これにより、その研究が同定されやすくなり、報告バイアスの減少にも役立つ。そのために登録番号・登録した際の研究名を明記する
29	プロトコル	正式な研究プロトコルを作成・公表し、いずれかでアクセス可能とすること。これにより、よりきめ細かな批判的吟味が可能となる。後付け解析を明確にし、研究の事前にプロトコルを公表しようという臨床試験や系統的レビューの最近の推奨と同じ流れである
30	資金源	研究者の科学的探究と倫理と相反する利益（企業の利潤等）があるか。財源を記載し明らかにすべきことは、診断研究においても例外ではない

(Bossuyt PM, et al. Radiology 2015; 277: 829. Copyright ©2015 Radiological Society of North America 和訳 酒井恭子・中山健夫)

・解析 Analysis

項目 18 必要とされたサンプルサイズとその見積り方法。

・考察 Discussion

項目 26・27 構造化された考察：研究の限界（バイアスの存在、統計学的不確実さ、一般化可能性など）、検査が対象とする集団と目的を考慮した示唆 (implication) を記述。

・その他の情報 Other information

項目 28 研究登録の詳細：新しい検査の確度の研究は臨床試験であり、研究開始前に登録が必要である。そうすることで、その研究に関心をもつ人々に見つけられやすくなり、研究者にとって都合の良い結果だけを発表する報告バイアスを減らすのにも役立つ。そのために登録番号・登録名を明記する。

項目 29 プロトコル Protocol：正式な研究プロトコルを作成・公表し、なんらかの方法でアクセス可能とすること。これにより、その研究に関心をもつ人々は、よりきめ細かな批判的吟味が可能となる。論文で報告された結果が著者にとって都合の良い後付け

解析であったかどうかを明確にするために、研究開始前にプロトコルを公表しようという臨床試験や系統的レビューの最近の潮流と同様である。

項目 30 資金源 source of funding：研究者の科学的探究の倫理と相反する利益（企業の利潤等）があるか、財源を記載して開示すべきであることは、診断研究も例外ではない。

Key となる用語 表として明示された（表 3）。

ま と め

診断確度の研究は、診断検査を評価する唯一の方法ではないが、診療現場での重要な役割を担うことは言うまでもない。残念ながら、その研究の質の担保は、一般的な介入・観察研究や系統的レビューに比して遅れている。STARD 声明 2003 が、診断確度研究の質を向上に役立ったか、高インパクトファクターの学術誌に掲載された論文を 2 つの時期（2000～2004 と 2012）で比較したところ、有意な向上がみられた⁷⁾。また、STARD 声明を順守している

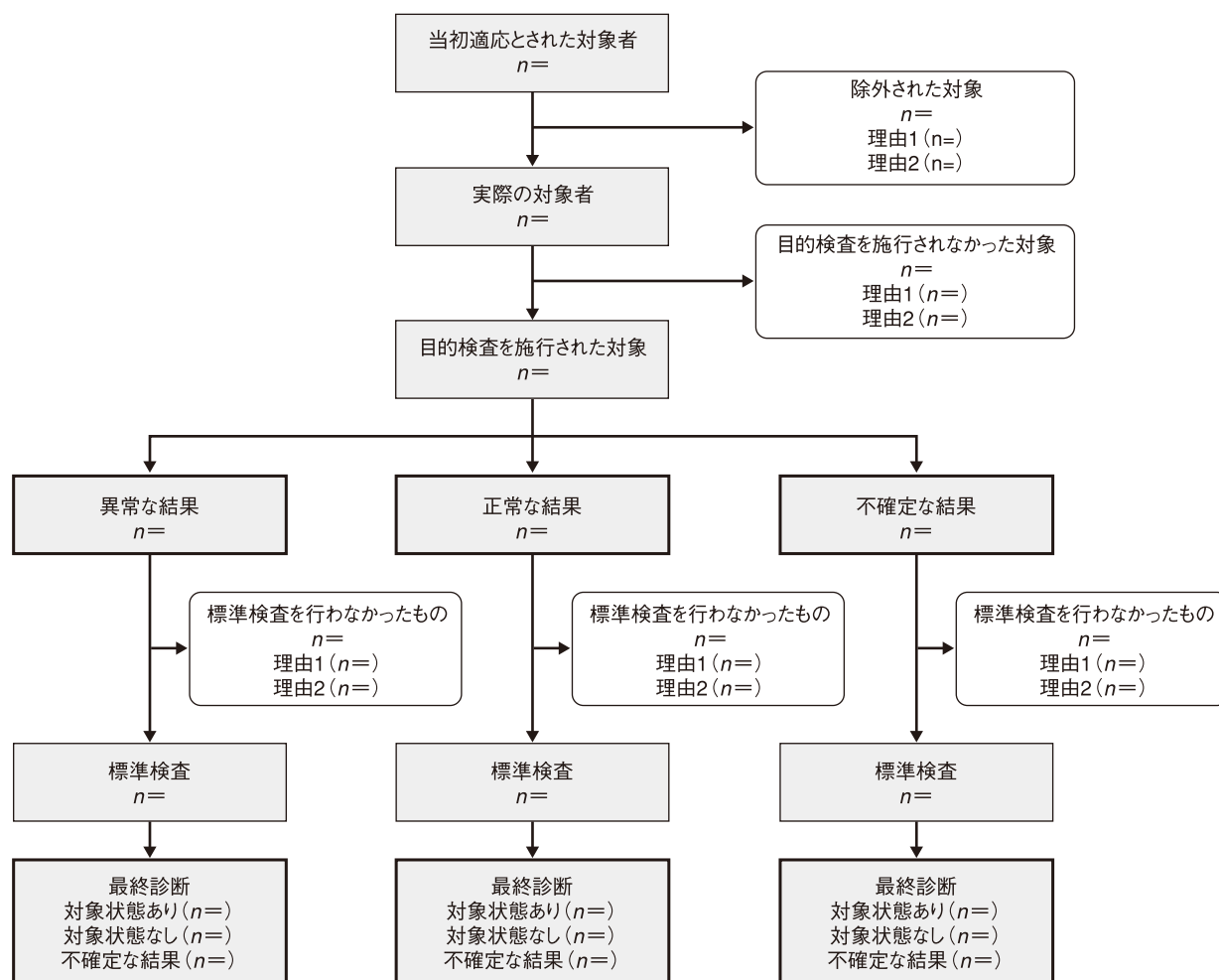


図 1 STARD2015 患者の募集

(Bossuyt PM, et al. Radiology 2015; 277: 830. Copyright ©2015 Radiological Society of North America 和訳 酒井恭子・中山健夫)

表 3 Key となる用語

用語 [Term]	説明 [Explanation]
検査 [Medical test]	患者の現在または将来における病態を把握するためのなんらかの検査
目的検査 [Index test]	検討対象となる検査
対象状態 [Target condition]	目的検査が指摘しようとする病気または臨床状態
臨床的標準検査 [Clinical reference standard]	対象状態の有無を示す最も理想的な標準検査。Gold standard はエラーがない検査であるべき
感受性 [Sensitivity]	目的検査の陽性者のうち対象状態ありの患者の割合
特異度 [Specificity]	目的検査の陰性者のうち対象状態なしの患者の割合
検査を行う目的 [Intended use of the test]	目的検査が、診断・スクリーニング・病気決定など、どのような状態の指摘を目的として用いられるか
検査の役割 [Role of the test]	目的検査が、その状態の他の検査に対し、どのような状況にあるか（たとえば、その検査との比較、置き換え、追加など）

(Bossuyt PM, et al. Radiology 2015; 277: 829. Copyright ©2015 Radiological Society of North America 和訳 酒井恭子・中山健夫)

論文は、その後の引用数が多かったという報告もある⁸⁾。以上のように、STARD 声明の順守が診断領域の論文の質を高める可能性があり、学術編集者には、ぜひ投稿規定で、STARD 声明 2015 に沿った論文の記述を投稿者に求め、査読者に対しても指針として活用できるように、積極的に STARD 声明 2015 の情報を提供されるように期待したい。

STARD 声明はまた、医学論文の執筆ガイドラインではあるが、論文の本体だけでなく、抄録やプロトコルの作成にも役立つであろう。STARD 声明 2015 は、EQUATOR の website で閲覧できる⁹⁾。研究と診療の質、そして患者にとっての益を向上させるため、本声明がさらに広く普及し、チェックリストがさまざまな環境で活用され、診断検査に関わる研究報告の質が向上することを願うものである。

文 献

- 1) Straus SE, Glasziou P, Richardson WS, et al. Evidence-based Medicine: How to practice and teach it. 4th ed. Elsevier; 2011.
- 2) Nelemans PJ, Leiner T, de Vet HCW, van Engelshoven JMA. Peripheral arterial disease: Meta-analysis of the diagnostic performance of MR angiography. Radiology 2000; 217: 105-14.
- 3) Devries SO, Hunink MGM, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. Acad Radiol 1996; 3: 361-9.
- 4) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. BMJ (Clinical research ed) 2003; 326: 41-4.
- 5) 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. ライフサイエンス出版; 2008.
- 6) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ (Clinical research ed) 2015; 351: h5527.
- 7) Korevaar DA, Wang J, van Enst WA, Leeftang MM, Hooft L, Smidt N, et al. Reporting diagnostic accuracy studies: some improvements after 10 years of STARD. Radiology 2015; 274: 781-9.
- 8) Dilauro M, McInnes MD, Korevaar DA, van der Pol CB, Petrich W, Walther S, et al. Is There an Association between STARD Statement Adherence and Citation Rate? Radiology 2016: 151384.
- 9) EQUATOR Website. <http://www.equator-network.org/> (2016/03/10 最終確認)