

STARD 2015: checklist aggiornata per il reporting degli studi di accuratezza diagnostica

Patrick M Bossuyt¹, Johannes B Reitsma², David E Bruns³, Constantine A Gatsonis⁴, Paul P Glasziou⁵, Les Irwig⁶, Jeroen G Lijmer⁷, David Moher^{8,9}, Drummond Rennie^{10,11}, Henrica C W de Vet¹², Herbert Y Kressel^{13,14}, Nader Rifai^{15,16}, Robert M Golub^{17,18}, Douglas G Altman¹⁹, Lotty Hooft²⁰, Daniël A Korevaar¹, Jérémie F Cohen^{1,21}, for the STARD Group

¹Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands; ²Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, University of Utrecht, Utrecht, the Netherlands; ³Department of Pathology, University of Virginia School of Medicine, Charlottesville, VA, USA; ⁴Center for Statistical Sciences, Brown University School of Public Health, Providence, RI, USA; ⁵Centre for Research in Evidence-Based Practice, Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Queensland, Australia; ⁶Screening and Diagnostic Test Evaluation Program, School of Public Health, University of Sydney, Sydney, New South Wales, Australia; ⁷Department of Psychiatry, Onze Lieve Vrouwe Gasthuis, Amsterdam, the Netherlands; ⁸Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada; ⁹School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, Ottawa, Canada; ¹⁰Peer Review Congress, Chicago, IL, USA; ¹¹Philip R Lee Institute for Health Policy Studies, University of California, San Francisco, CA, USA; ¹²Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands; ¹³Department of Radiology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; ¹⁴Radiology Editorial Office, Boston, MA, USA; ¹⁵Department of Laboratory Medicine, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA; ¹⁶Clinical Chemistry Editorial Office, Washington, DC, USA; ¹⁷Division of General Internal Medicine and Geriatrics and Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; ¹⁸JAMA Editorial Office, Chicago, IL, USA; ¹⁹Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK; ²⁰Dutch Cochrane Centre, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, University of Utrecht, Utrecht, the Netherlands; ²¹INSERM UMR 1153 and Department of Pediatrics, Necker Hospital, AP-HP, Paris Descartes University, Paris, France.

ABSTRACT

Una delle principali fonti evitabili di spreco nella ricerca biomedica è costituita dal reporting incompleto degli studi che spesso non forniscono le informazioni necessarie per la loro identificazione, valutazione critica e riproducibilità. Per migliorare la qualità dei report degli studi di accuratezza diagnostica è stato sviluppato lo statement sugli standard per il reporting degli studi di accuratezza diagnostica: *Standards for Reporting Diagnostic Accuracy* (STARD). Questo articolo presenta la versione 2015 di STARD, lista aggiornata di 30 item essenziali che

dovrebbero essere inseriti in tutti i report degli studi di accuratezza diagnostica. L'aggiornamento comprende recenti evidenze su fonti di bias e variabilità nell'accuratezza diagnostica e ha lo scopo di facilitare l'utilizzo di STARD. STARD 2015 può contribuire a migliorare completezza e trasparenza nel reporting degli studi di accuratezza diagnostica.

Citazione. Bossuyt PM, Reitsma JB, Bruns DE et al. STARD 2015: aggiornamento della checklist per il reporting degli studi di accuratezza diagnostica. *Evidence* 2016;8(1): e1000129.

Pubblicato 26 gennaio 2016

Copyright. 2016 Bossuyt et al. Questo è un articolo open-access, distribuito con licenza *Creative Commons Attribution*, che ne consente l'utilizzo, la distribuzione e la riproduzione su qualsiasi supporto esclusivamente per fini non commerciali, a condizione di riportare sempre autore e citazione originale.

Fonti di finanziamento. Nessun finanziamento esplicito per lo sviluppo di STARD 2015. L'Academic Medical Center dell'Università di Amsterdam (Paesi Bassi) ha finanziato in parte il meeting del comitato direttivo STARD senza influenzare sviluppo e disseminazione della checklist. I membri del comitato direttivo STARD e i membri del gruppo STARD hanno sostenuto personalmente i costi.

Conflitti di interesse. Tutti gli autori hanno sottoscritto la disclosure per i conflitti di interesse prevista dalla rivista *Clinical Chemistry*: Rifai N lavora per *Clinical Chemistry*, AACCC; Gatsonis CA è membro del RSNA Research Development Committee.

Provenienza. Tradotto da: Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;351:h5527

Questo articolo è stato pubblicato contemporaneamente nell'ottobre 2015 su *BMJ*, *Radiology* e *Clinical Chemistry* con licenza Creative Commons CC BY-NC: <http://creativecommons.org/licenses/by-nc/4.0>.

*E-mail: p.m.bossuyt@amc.uva.nl

I ricercatori discutono e scrivono dei propri studi, non solo perchè soddisfatti o meno dei risultati ottenuti, ma anche per consentire di comprenderne la validità dei metodi, permettere ai colleghi di replicarli e per rendere noti i risultati a medici, altri professionisti sanitari e decisori, le cui decisioni vengono guidate da evidenze metodologicamente robuste.

Purtroppo in molte aree della medicina sono state evidenziate carenze nel reporting della ricerca¹. Elementi essenziali dei metodi vengono spesso descritti in maniera inadeguata e talvolta del tutto omessi, rendendo difficile se non impossibile sia la valutazione critica che la possibilità di ripetere lo studio. I risultati vengono a volte riportati in modo selettivo, oppure interpretati con ingiustificato ottimismo²⁻⁴. Questi comportamenti riducono il valore degli studi primari e condizionano negativamente quelli secondari, quali revisioni sistematiche e linee guida.

I report degli studi che valutano l'accuratezza dei test diagnostici non costituiscono eccezione: infatti, un numero crescente di valutazioni ha identificato numerose carenze nel loro reporting⁵. Si tratta di studi dove un test diagnostico viene valutato rispetto ad uno standard di riferimento clinico, o gold standard; i risultati sono in genere riportati come stime di sensibilità e specificità del test, che esprimono l'accuratezza del test nell'identificare correttamente i pazienti con la condizione target. Altre misure utilizzate sono la curva ROC (*Receiver Operating Characteristics*) e i valori predittivi, positivo e negativo.

Nonostante la loro apparente semplicità, questi studi sono a rischio di bias^{6,7}. Ad esempio, se nell'analisi finale non vengono inclusi tutti i pazienti sottoposti al test in studio, oppure se vengono inclusi solo i controlli sani, le stime di accuratezza del test possono non riflettere la sua reale performance nelle applicazioni cliniche. Eppure spesso i report degli studi non contengono queste informazioni fondamentali.

È oggi ben documentato che sensibilità e specificità non sono caratteristiche fisse di un test diagnostico: infatti, i risultati falsi positivi e falsi negativi variano in relazione al contesto, alle modalità di selezione dei pazienti e ai test diagnostici a cui sono già stati sottoposti. Purtroppo molti autori non forniscono nel report dello studio adeguati dettagli su contesto clinico, tempo, luogo e modalità con cui sono stati identificati e reclutati i partecipanti eleggibili⁸. Inoltre, le stime di sensibilità e specificità possono differire a causa della variabilità delle definizioni dello standard di riferimento con cui viene confrontato il test in studio. Pertanto, queste informazioni dovrebbero essere disponibili nel report dello studio.

Lo STARD Statement 2003

Per migliorare completezza e trasparenza dei report sugli studi di accuratezza diagnostica, un gruppo di ricercatori, direttori di riviste e altri stakeholders ha sviluppato una li-

sta di item fondamentali che dovrebbero essere inclusi nel report di ogni studio. Il principio guida per lo sviluppo della checklist è stato quello di selezionare gli elementi che, se descritti, aiuterebbero il lettore a giudicare il rischio di bias dello studio e valutare l'applicabilità dei risultati e la validità di conclusioni e raccomandazioni degli autori.

Nel 2003 è stato pubblicato in 24 riviste scientifiche lo statement relativo agli standard per il reporting degli studi di accuratezza diagnostica – *Standards for Reporting Diagnostic Accuracy* (STARD)⁹. Lo statement è stato accompagnato da editoriali e commenti in numerose pubblicazioni e sostenuto da molte altre riviste. Dopo la pubblicazione dello STARD alcune valutazioni hanno evidenziato miglioramenti piccoli, ma statisticamente significativi, nel reporting degli studi di accuratezza diagnostica [guadagno medio 1.4 item (IC 95% 0.7-2.2)]^{5,10}. Progressivamente i report degli studi hanno iniziato ad includere un numero maggiore di item della checklist, ma la situazione è ancora lontana da quella ottimale.

Metodi per lo sviluppo dello STARD 2015

Nel 2013 il comitato direttivo STARD, che rivede periodicamente la letteratura per selezionare studi potenzialmente rilevanti ha deciso che era il momento di aggiornare la checklist con due obiettivi principali: 1. integrare le recenti evidenze su fonti di bias, problemi di applicabilità e identificazione dei fattori che inducono a interpretazioni "generose" degli studi di accuratezza dei test diagnostici; 2. semplificare l'uso della checklist. L'aggiornamento ha tenuto in considerazione la necessità di armonizzazione STARD con altre linee guida per il reporting, quali il CONSORT 2010¹¹.

Una descrizione completa del processo di aggiornamento e del rationale delle modifiche sono disponibili sul sito EQUATOR - *Enhancing the Quality and Transparency of Health Research* (www.equator-network.org/reporting-guidelines/stard). In sintesi, i membri del gruppo STARD 2003 sono stati invitati a partecipare al processo di aggiornamento, a identificare nuovi membri e a commentare gli obiettivi generali dell'aggiornamento. Il gruppo STARD oggi ha raggiunto 85 membri che comprendono ricercatori, direttori di riviste, giornalisti, esperti nella sintesi delle evidenze, finanziatori e altri stakeholders.

I membri del gruppo STARD hanno suggerito e, quindi, approvato le proposte di aggiornamento attraverso una survey via web in due round. Questo ha permesso di definire una bozza della lista degli item principali, discussa dal comitato direttivo in un meeting di due giorni tenutosi ad Amsterdam nel settembre 2014. La lista è stata quindi sottoposta a un test pilota da parte di ricercatori junior e senior, revisori e direttori di riviste.

La struttura generale di STARD 2015 è simile a quella di STARD 2003. La checklist (tabella 1) riporta i 30 item,

Tabella 1. La checklist STARD 2015*

SEZIONE E TOPIC	N° ITEM	DESCRIZIONE
Titolo o abstract		
	1	Identificare l'articolo come uno studio di accuratezza diagnostica utilizzando almeno una misura di accuratezza (es. sensibilità, specificità, valori predittivi o area sotto la curva ROC)
Abstract		
	2	Abstract strutturato che comprende: disegno dello studio, metodi, risultati e conclusioni (per una guida specifica fare riferimento al documento STARD per gli abstract)
Introduzione		
	3	Background scientifico e clinico, compreso l'uso previsto e il ruolo clinico del test in studio
	4	Obiettivi e ipotesi dello studio
Metodi		
Disegno dello studio	5	Specificare se la raccolta dati è stata pianificata prima di eseguire il test in studio e lo standard di riferimento (studio prospettico), oppure dopo (studio retrospettivo)
Partecipanti	6	Criteri di eleggibilità
	7	Modalità per identificare i partecipanti potenzialmente eleggibili (es. sulla base dei sintomi, di risultati di test precedenti, del loro inserimento in un registro)
	8	Dove e quando sono stati identificati i partecipanti potenzialmente eleggibili: setting, luogo e date
	9	Specificare se i partecipanti sono stati arruolati in maniera consecutiva, casuale o opportunistica
Metodi del test	10a	Test in studio, descritto con sufficienti dettagli da renderlo riproducibile
	10b	Standard di riferimento, descritto con sufficienti dettagli da renderlo riproducibile
	11	Razionale per la scelta dello standard di riferimento (nel caso in cui esistano alternative)
	12a	Definizione e rationale dei cut-off di positività o delle categorie di risultati per il test in studio, distinguendo quelli pre-definiti da quelli esplorativi
	12b	Definizione e rationale dei cut-off di positività o delle categorie di risultati per lo standard di riferimento, distinguendo quelli pre-definiti da quelle esplorativi
	13a	Specificare se le informazioni cliniche e i risultati dello standard di riferimento erano disponibili a chi ha eseguito o refertato il test in studio
	13b	Specificare se le informazioni cliniche e i risultati del test in studio erano disponibili a chi ha valutato lo standard di riferimento
Analisi	14	Metodi per valutare o confrontare le misure di accuratezza diagnostica
	15	Modalità per gestire i risultati indeterminati del test in studio e/o dello standard di riferimento
	16	Modalità per gestire i dati mancanti del test in studio e/o dello standard di riferimento
	17	Eventuali analisi di variabilità dell'accuratezza diagnostica, distinguendo quelle pre-definite da quelle esplorative
	18	Dimensioni del campione previsto e modalità per determinarlo
Risultati		
Partecipanti	19	Flusso dei partecipanti, utilizzando un diagramma
	20	Caratteristiche demografiche e cliniche basali dei partecipanti
	21a	Spettro di gravità della malattia nei partecipanti che presentano la condizione target
	21b	Distribuzione delle diagnosi alternative nei partecipanti che non presentano la condizione target
	22	Riportare l'intervallo temporale e tutti gli interventi clinici effettuati tra l'esecuzione del test in studio e lo standard di riferimento
Risultati del test	23	Riportare in una tabella 2 x 2 i risultati del test in studio (o loro distribuzione) e dello standard di riferimento
	24	Stime di accuratezza diagnostica e loro precisione (es. intervalli di confidenza al 95%)
	25	Riportare tutti gli eventi avversi conseguenti all'esecuzione del test in studio o dello standard di riferimento
Discussione		
	26	Limiti dello studio, comprese le fonti di potenziale bias, l'incertezza statistica e la generalizzabilità
	27	Implicazioni per la pratica, compreso l'utilizzo previsto e il ruolo clinico del test in studio
Altre informazioni		
	28	Numero di registrazione dello studio e nome del registro
	29	Indicare dove è possibile reperire il protocollo integrale dello studio
	30	Fonti di finanziamento e altre forme di supporto; ruolo dei finanziatori

*All'inizio della riga di ciascun item gli autori dovrebbero indicare il numero di pagina del manoscritto dove viene riportato l'item

raggruppati in sezioni coerenti con la struttura IMRaD degli articoli scientifici (Introduzione, Metodi, Risultati, Discussione). Numerosi item sono identici a quelli della versione 2003, mentre altri sono stati riformulati, combinati o suddivisi se complessi. La tabella 2 descrive i nuovi item e la tabella 3 elenca le parole chiave. Un diagramma per descrivere il flusso di partecipanti è ora previsto in tutti i report (figura).

Scopo

STARD 2015 sostituisce integralmente la versione del 2003: pertanto chi vuole fare riferimento a STARD è invitato a citare questo articolo. La checklist rappresenta un

set minimo di item, ma il report di uno studio generalmente richiede un numero maggiore di informazioni. In ogni caso, un report adeguato di uno studio di accuratezza diagnostica deve comprendere tutti gli item applicabili.

Gli autori sono invitati ad utilizzare STARD nella stesura dei report dei loro studi. I revisori possono utilizzare la checklist per verificare che il manoscritto riporti tutte le informazioni essenziali e suggerire modifiche se mancano item fondamentali.

Le riviste che hanno supportato lo STARD sin dal 2003 dovrebbero raccomandare l'utilizzo di questa versione aggiornata e incoraggiare l'aderenza dei manoscritti, sperando che anche altre riviste ed editori promuovano STARD

Tabella 2. Sintesi dei nuovi item della checklist STARD 2015

N.	Item	Razionale
2	Abstract strutturato	Gli abstract sono sempre più utilizzati per identificare gli elementi fondamentali del disegno dello studio e dei risultati
3	Uso previsto e ruolo clinico del test in studio	La descrizione delle applicazioni del test permette ai lettori di interpretare le implicazioni pratiche delle stime di accuratezza diagnostica riportate
4	Ipotesi dello studio	La mancanza di ipotesi specifiche può favorire l'interpretazione ottimistica dei risultati e la distorsione delle conclusioni
18	Dimensioni del campione	I lettori hanno la necessità di valutare la precisione stimata e la potenza statistica dello studio e se gli autori sono riusciti a reclutare il numero previsto di partecipanti
26-27	Discussione strutturata	Per evitare di trarre conclusioni non dimostrate, gli autori sono invitati a riportare i limiti dello studio e a formulare le conclusioni sulla base delle applicazioni dei test valutati (item 3)
28	Registrazione	Gli studi prospettici per valutare l'accuratezza di un test sono trial che possono essere registrati prima del loro inizio in registri di trial clinici (es. ClinicalTrials.gov), per facilitarne l'identificazione e prevenire il reporting selettivo
29	Protocollo	Il protocollo integrale dello studio, che contiene maggiori informazioni sui metodi, può essere reso disponibile per una valutazione critica più dettagliata
30	Fonti di finanziamento	Occorre essere consapevoli degli effetti potenzialmente dannosi dei conflitti di interesse tra obblighi etici e scientifici dei ricercatori e altri obiettivi come quelli economici; gli studi di accuratezza diagnostica non rappresentano una eccezione alla regola

Tabella 3. Glossario

Termine	Spiegazione
Test clinico	Qualunque metodo utilizzato per raccogliere informazioni aggiuntive sullo stato di salute attuale e futuro del paziente
Test in studio	Il test valutato dallo studio di accuratezza diagnostica (<i>index test</i>)
Condizione target	La malattia o condizione clinica che il test in studio deve diagnosticare
Standard di riferimento	Il metodo più accurato disponibile per stabilire la presenza/assenza della condizione target: il gold standard identifica uno standard di riferimento ideale esente da errori
Sensibilità	Percentuale di soggetti affetti dalla condizione target e positivi al test in studio
Specificità	Percentuale di soggetti senza la condizione target e negativi al test in studio
Uso previsto del test	Uso del test in studio per diagnosi, screening, stadiazione, monitoraggio, sorveglianza, prognosi o altri obiettivi
Ruolo del test	Ruolo del test in studio rispetto ad altri test usati per la stessa condizione (es. nel triage, in sostituzione o in aggiunta a un altro test, nuovo test)

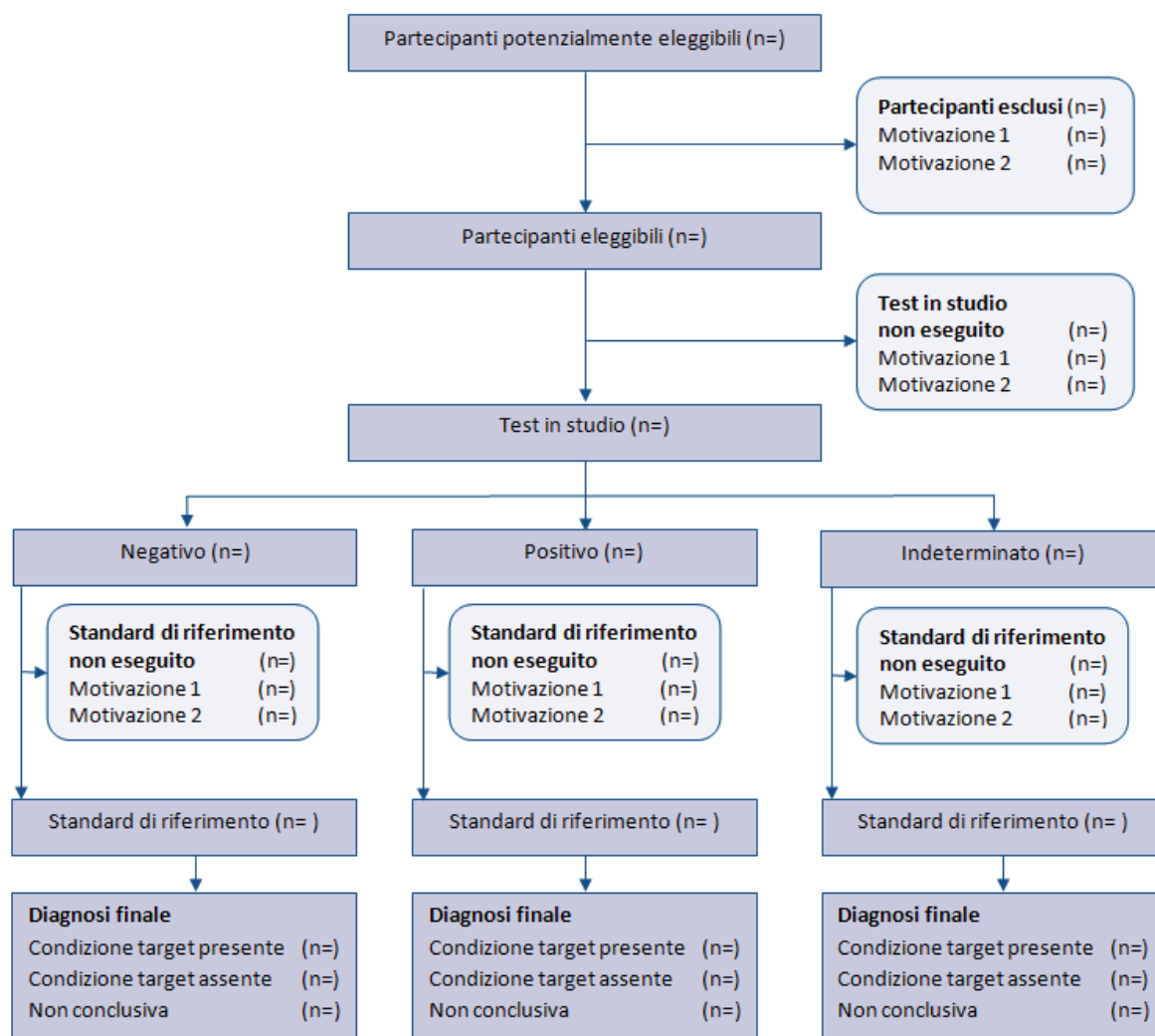


Figura. Diagramma di flusso STARD

e altre linee guida per il reporting. Finanziatori e istituti di ricerca possono promuovere o rendere obbligatoria l'aderenza a STARD per massimizzare il *value* della ricerca diagnostica e i prodotti di letteratura secondaria correlati.

STARD può essere utile anche per il reporting di altri studi che valutano la performance dei test diagnostici: studi prognostici, che possono classificare i pazienti in relazione a un evento futuro; studi di monitoraggio, nei quali si suppone che i test diagnostici individuino o siano predittivi di un evento avverso o della mancata risposta terapeutica, studi di valutazione di marcatori per la scelta della terapia, etc. Tuttavia, se la maggior parte degli item dello STARD può essere utilizzata per il reporting di questi studi, la checklist è destinata principalmente agli studi di accuratezza diagnostica.

L'accuratezza diagnostica non è l'unica misura della performance di un test, né necessariamente la più significativa¹². Ad esempio, l'accuratezza incrementale legata

alla combinazione di test può essere più informativa di quella relativa a un singolo test¹³. Per i test misurati con variabili continue, la dicotomizzazione positivo/negativo non è sempre indicata; in questi casi, i metodi di calcolo e grafici più adeguati per esprimere la performance del test sono differenti, anche se molte delle precauzioni metodologiche potrebbero essere le stesse e STARD può migliorare l'informatività del report dello studio. Altre linee guida per il reporting individuano categorie di test più specifiche, come TRIPOD (*Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis*) per i modelli di previsione multivariati¹⁴.

Anche se STARD si focalizza sul reporting integrale degli studi di accuratezza diagnostica, i suoi item possono essere utili per la stesura di protocolli di studio, di abstract destinati a conferenze e per le informazioni da inserire nei registri di trial. Ulteriori iniziative sono in corso per fornire indicazioni più specifiche per ciascuna di queste applicazioni.

Estensioni e applicazioni di STARD 2015

STARD 2015 è stato progettato per essere applicato a tutti i test diagnostici. Al fine di una maggiore diffusione e accettazione da parte di autori, revisori e direttori di riviste, il gruppo STARD ha preferito una checklist unica per tutti gli studi di accuratezza diagnostica, rispetto a checklist differenti per tipologie diverse di test (imaging, biochimica, istopatologia).

Ovviamente, una checklist generica può richiedere ulteriori istruzioni per il reporting, con maggiori informazioni rispetto a specifici test, specifiche applicazioni o tipologie di analisi. Ad esempio, queste indicazioni potrebbero descrivere i metodi preferiti per lo studio e il reporting della misura di incertezza, senza modificare gli altri item della checklist. Il gruppo STARD approva lo sviluppo di queste estensioni e invita i gruppi interessati a contattare il comitato direttivo prima di svilupparli.

Analogamente, è possibile sviluppare ulteriori linee guida per facilitare l'uso di STARD in specifiche applicazioni, quali la raccolta dell'anamnesi e dell'esame obiettivo¹⁵, oppure in specifiche condizioni target quali la demenza¹⁶.

Disponibilità

La checklist STARD 2015 e tutti i documenti correlati sono disponibili sul sito web di EQUATOR, iniziativa internazionale finalizzata ad aumentare il *value* della ricerca, promuovendo un reporting trasparente e completo e diffondendo l'utilizzo di rigorose linee guida per il reporting^{17,18}. Il gruppo STARD ritiene che una più stretta collaborazione con EQUATOR ed altri autori di linee guida per il reporting faciliterà il raggiungimento di obiettivi condivisi. Sul sito web di EQUATOR è disponibile anche l'aggiornamento del documento di spiegazione ed elaborazione del 2003, che fornisce per ciascun item rationale ed esempi.

La checklist STARD è pubblicata con licenza *Creative Commons* che permette a tutti di utilizzare e distribuire il lavoro, citando la fonte originale. La versione originale dello STARD 2015 è in lingua inglese, ma numerosi gruppi hanno già realizzato traduzioni in altre lingue. Il gruppo STARD approva queste traduzioni, preferibilmente realizzate da gruppi di ricercatori attraverso un processo di sviluppo ciclico, con back-translation in lingua originale e test pilota¹⁹. STARD è un marchio registrato e il comitato direttivo mantiene il diritto esclusivo di utilizzare il termine "STARD" per identificare prodotti o servizi.

Aumentare il *value*, ridurre gli sprechi

Il comitato direttivo è consapevole che una lista di item essenziali non è sufficiente per ottenere miglioramenti sostanziali nella completezza del reporting, come dimostrano i modesti risultati ottenuti dopo l'introduzione della prima versione della checklist nel 2003. Di fatto, la checklist rappresenta il punto di partenza per costruire

strumenti più specifici, basati sulla lista di item essenziali dello STARD, finalizzati a promuovere un reporting più completo e trasparente: una lista di controllo e uno strumento per supportare gli autori nella stesura del manoscritto, strumenti per revisori e direttori di riviste, video e materiali didattici.

Il reporting incompleto è stato identificato come una delle fonti di spreco evitabile nella ricerca biomedica¹. Dalla nascita di STARD, numerose iniziative sono state intraprese per migliorare la riproducibilità della ricerca e promuoverne una maggiore trasparenza²⁰. L'augurio è che l'aggiornamento e l'implementazione dello STARD, aiuti autori, direttori di riviste, revisori, lettori e decisori a raccogliere, valutare e applicare le evidenze necessarie a rafforzare decisioni e raccomandazioni sui test diagnostici. Tutti possono trarre beneficio da un reporting più informativo e trasparente della ricerca: ricercatori, professionisti sanitari, cittadini e pazienti.

COLLABORATORI DEL GRUPPO STARD

Todd Alonzo, Douglas G Altman, Augusto Azuara-Blanco, Lucas Bachmann, Jeffrey Blume, Patrick M Bossuyt, Isabelle Boutron, David Bruns, Harry Büller, Frank Buntinx, Sarah Byron, Stephanie Chang, Jérémie F Cohen, Richelle Cooper, Joris de Groot, Henrica C W de Vet, Jon Deeks, Nandini Dendukuri, Jac Dinnes, Kenneth Fleming, Constantine A Gatsonis, Paul P Glasziou, Robert M Golub, Gordon Guyatt, Carl Heneghan, Jørgen Hilden, Lotty Hooft, Rita Horvath, Myriam Hunink, Chris Hyde, John Ioannidis, Les Irwig, Holly Janes, Jos Kleijnen, André Knottnerus, Daniël A Korevaar, Herbert Y Kressel, Stefan Lange, Mariska Leeflang, Jeroen G Lijmer, Sally Lord, Blanca Lumberras, Petra Macaskill, Erik Magid, Susan Mallett, Matthew McInnes, Barbara McNeil, Matthew McQueen, David Moher, Karel Moons, Katie Morris, Reem Mustafa, Nancy Obuchowski, Eleanor Ochodo, Andrew Onderdonk, John Overbeke, Nitika Pai, Rosanna Peeling, Margaret Pepe, Steffen Petersen, Christopher Price, Philippe Ravaut, Johannes B. Reitsma, Drummond Rennie, Nader Rifai, Anne Rutjes, Holger Schunemann, David Simel, Iveta Simera, Nynke Smidt, Ewout Steyerberg, Sharon Straus, William Summerskill, Yemisi Takwoingi, Matthew Thompson, Ann van de Bruel, Hans van Maanen, Andrew Vickers, Gianni Virgili, Stephen Walter.

MATERIALE SUPPLEMENTARE

Checklist STARD 2015
Diagramma di flusso

CONTRIBUTO DEGLI AUTORI

Tutti gli autori confermano di aver contribuito al contenuto intellettuale di questo articolo e soddisfano i seguenti requisiti: (a) contributi significativi all'ideazione e

al disegno, acquisizione dei dati o analisi e interpretazione dei dati; (b) stesura o revisione dei contenuti intellettuali dell'articolo; e (c) approvazione finale dell'articolo pubblicato.

NOTE ALLA VERSIONE ITALIANA

La Fondazione GIMBE ha realizzato la traduzione italiana dell'articolo senza alcun finanziamento istituzionale o commerciale.

TEAM CHE HA REALIZZATO LA VERSIONE ITALIANA

Responsabili scientifici

Antonino Cartabellotta, Fondazione GIMBE
Romolo M. Dorizzi, Azienda USL Romagna

Coordinamento editoriale

Marco Mosti, Fondazione GIMBE

Traduzione

Elisa Esposito, Palermo
Federica Riccio, Specialista in Igiene e Medicina Preventiva. Casale Monferrato (AL)

Revisione editoriale

Elena Cottafava, Fondazione GIMBE
Roberto Luceri, Fondazione GIMBE

BIBLIOGRAFIA

1. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383:267-76.
2. Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA* 2010;303:2058-64.
3. Ochodo EA, de Haan MC, Reitsma JB, et al. Overinterpretation and misreporting of diagnostic accuracy studies: evidence of "spin." *Radiology* 2013;267:581-8.
4. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;302:977-84.
5. Korevaar DA, Wang J, van Enst WA, et al. Reporting diagnostic accuracy studies: some improvements after 10 years of STARD. *Radiology* 2015;274:781-9.
6. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061-6.
7. Whiting PF, Rutjes AW, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. *J Clin Epidemiol* 2013;66:1093-104.
8. Irwig L, Bossuyt P, Glasziou P, Gatsonis C, Lijmer J. Designing studies to ensure that estimates of test accuracy are transferable. *BMJ* 2002;324:669-71.
9. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Radiology* 2003;226:24-8.
10. Korevaar DA, van Enst WA, Spijker R, Bossuyt PM, Hooft L.

Reporting quality of diagnostic accuracy studies: a systematic review and meta-analysis of investigations on adherence to STARD. *Evid Based Med* 2014;19:47-54.

11. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol* 2010;63:834-40.
12. Bossuyt PM, Reitsma JB, Linnert K, Moons KG. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. *Clin Chem* 2012;58:1636-43.
13. Moons KG, de Groot JA, Linnert K, Reitsma JB, Bossuyt PM. Quantifying the added value of a diagnostic test or marker. *Clin Chem* 2012;58:1408-17.
14. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594.
15. Simel DL, Rennie D, Bossuyt PM. The STARD statement for reporting diagnostic accuracy studies: application to the history and physical examination. *J Gen Intern Med* 2008;23:768-74.
16. Noel-Storr AH, McCleery JM, Richard E, et al. Reporting standards for studies of diagnostic test accuracy in dementia: the STARDdem Initiative. *Neurology* 2014;83:364-73.
17. Altman DG, Simera I, Hoey J, Moher D, Schulz K. EQUATOR: reporting guidelines for health research. *Lancet* 2008;371:1149-50.
18. Simera I, Moher D, Hirst A, et al. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med* 2010;8:24.
19. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186-91.
20. Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature* 2014;505:612-3.