ABSTRACT 2 Structured summ (for specific guid (for specific guid (for specific guid for specific guid for specific guid (for specific guid for specific and clir (for specific and clir 4 Study objectives) METHODS Study design 5 Whether data converse performed were performed for such as symptor (such as symptor (such as symptor such as symptor such as symptor for such as symptor (such as symptor such as symptor for such as symptor such as symptor for such as symptor for such as symptor such as symptor for such as sympt			
ABSTRACT 2 Structured summ (for specific guid introduction) 3 Scientific and clir 4 Study objectives METHODS Study design 5 Whether data come were performed year performed. Participants 6 Eligibility criteria 7 On what basis portion (such as symptor such as symptor and performed in the performed in the summer of the index test, in suff of the index test, in suff of the index test, in suff of the index test, in the index te			
ABSTRACT 2 Structured summ (for specific guid introduction) 3 Scientific and clir 4 Study objectives METHODS Study design 5 Whether data come were performed year performed. Participants 6 Eligibility criteria 7 On what basis portion (such as symptor such as symptor) 8 Where and when 9 Whether particip 11 Rationale for chome 12 Definition of and of the index test, in suff of the reference 12 Definition of and of the reference 13 Whether clinical to the performer 13 Whether clinical to the assessors of the assessors of the index test, in suff of the reference 15 How indetermina 16 How missing data to the assessors of 17 Any analyses of whether clinical to the performer 18 Intended sample 19 Flow of participants 19 Flow of participants 20 Baseline demogration of the results of 21 Time interval and 22 Time interval and 22 Time interval and 23 Cross tabulation by the results of 24 Estimates of diagrams 25 Any adverse ever 25 DISCUSSION	study of diagnostic accuracy using at least one measure of		
INTRODUCTION 3 Scientific and clir 4 Study objectives METHODS Study design 5 Whether data come were performed Participants 6 Eligibility criteria 7 On what basis porticular symptom 8 Where and when 9 Whether particip 10b Reference standa 11 Rationale for chome standa 12 Definition of and soft the index test, so the performer standa standard to the performer standard to the performer standard to the sassessors of the index test, so the sassessors of the sassessors of the index test, so the sassessors of the index test, so the sassessors of the index test, so the sassessors of the sassessors of the index test, so the sass	sensitivity, specificity, predictive values, or AUC)	\checkmark	1
INTRODUCTION 3 Scientific and clir 4 Study objectives METHODS Study design 5 Whether data come were performed were performed were performed. Participants 6 Eligibility criteria 7 On what basis possuch as symptor 8 Whether participal index test, in suff lob Reference standard of the index test, in suff of the inde			
INTRODUCTION 3 Scientific and clir 4 Study objectives METHODS Study design 5 Whether data cowere performed Participants 6 Eligibility criteria 7 On what basis porticipate and when such as symptor 8 Where and when such as symptor 100 Reference standar and of the index test, in suff and of the index test, in a suff and of the index test, in suff and of the index test,	ary of study design, methods, results, and conclusions		······
METHODS Study design Farticipants 6 Eligibility criteria 7 On what basis por (such as symptor) 8 Where and when 9 Whether participal 10 Reference standard of the index test, in suff of the index test, in 20 Definition of and of the reference 12a Whether clinical to the performer 13b Whether clinical to the assessors of the assessors of the index test, in 3d Whether clinical to the assessors of the index test, in 5d How missing data to the assessors of the index test, in 5d How missing data to the assessors of the index test, in 5d How missing data to the assessors of the index test, in 5d How missing data to the assessors of the index test, in 5d How missing data to the assessors of the index test, in 5d How missing data to the assessors of the index test, in 5d How indeterminate 15d How missing data to the assessors of the index test, in 5d How indeterminate 15d How missing data to the assessors of the index test, in 5d How indeterminate 15d How missing data to the assessors of the index test, in 5d How indeterminate 15d How missing data to the assessors of the index test, in 5d How indeterminate 15d How missing data to the assessors of the index test, in 5d How indeterminate 15d How missing data to the assessors of the index test, in 5d How indeterminate 15d How indeterm	ance, see STARD for Abstracts)	V	3
METHODS Study design 5 Whether data comerce performed Participants 6 Eligibility criteria 7 On what basis porticipants (such as symptor) 8 Where and when 9 Whether participants (10b Reference standard) 11 Rationale for chomer (12a Definition of and of the index test, and of the index test, and of the reference (13a Whether clinical to the performer (13b Whether clinical to the assessors of the desired of the index test) 15 How indeterminates (15 How missing data (17 Any analyses of whether clinical to the desired of the index test) 18 Intended sample (17 Any analyses of whether clinical to the desired of the index test) 19 Flow of participants (19 Elow of participants) 20 Baseline demogration of second of the index test) 21a Distribution of all the index test (17 Any analyses of whether clinical to the assessors of the index test) 18 Intended sample (18 Intended sample) RESULTS Participants (19 Elow of participants) 21a Distribution of all the index test (18 Intended sample) 22 Time interval and (18 Intended sample) 23 Cross tabulation by the results of the stimulation of the second of t		. /	
METHODS Study design 5 Whether data cowere performed Participants 6 Eligibility criteria 7 On what basis porticular and when such as sympton 8 Where and when such as sympton 10 Index test, in sufficial and for chower and such as sympton 10 Reference stands 11 Rationale for chower and such as sympton and such as sympton 12a Definition of and such as sympton and such as sympton 12b Definition of and such as sympton and such as sessors of the index test, and such as sympton and such as sympton 13a Whether clinical to the performer 13b Whether clinical to the assessors of the index sympton and sympton	ical background, including the intended use and clinical role of the inde	ex test	4-5
Study design Participants 6 Eligibility criteria 7 On what basis por (such as sympton) 8 Where and when on the participants 10 Index test, in suff (such as sympton) 11 Rationale for chood (such as sympton) 12 Definition of and of the index test, in suff (such as seen as a sympton) 12 Definition of and of the index test, in suff (such as a sympton) 13 Whether clinical to the performen and of the reference sympton (such as a sympton) 13 Whether clinical to the performen and to the assessors of the sympton (such as a sympton) 14 Methods for estimates and the sympton (such as a sympton) 15 How indeterminates and sympton (such as a sympton) 16 How missing data and sympton (such as a sympton) 17 Any analyses of voch as a sympton (such as a sympton) 18 Intended sample and sympton (such as a sympton) 20 Baseline demogration of all sympton (such as a sympton) 21 Distribution of all condition of the results of sympton (such as a sympton) 22 Time interval and condition of the results of sympton (such as a sympton) 23 Cross tabulation by the results of sympton (such as a sympton) 24 Estimates of diag condition (such as a sympton) 25 Any adverse ever	and hypotheses	\	5
were performed Participants 6 Eligibility criteria 7 On what basis por (such as symptor) 8 Where and when 9 Whether particip Test methods 10a Index test, in suff 10b Reference standard 11 Rationale for cho 12a Definition of and of the index test, 12b Definition of and of the reference 13a Whether clinical to the performer 13b Whether clinical to the assessors of the assessors of the index test, 15 How indeterminary 16 How missing datary 17 Any analyses of variation 18 Intended sample 19 Plow of participants 19 Plow of participants 20 Baseline demogration 21a Distribution of all 22 Time interval and 22 Time interval and 23 Cross tabulation by the results of 24 Estimates of diag 25 Any adverse ever 21b DISCUSSION		¥	
Participants 6 Eligibility criteria 7 On what basis por (such as symptor) 8 Where and when 9 Whether particip Test methods 10a Index test, in suff 10b Reference standar 11 Rationale for chood 12a Definition of and of the index test, 12b Definition of and of the reference 13a Whether clinical to the performer 13b Whether clinical to the assessors of the Analysis 14 Methods for estimate 15 How indeterminate 16 How missing data 17 Any analyses of wind 18 Intended sample 17 Any analyses of wind 18 Intended sample 19 Plow of participants 20 Baseline demogration 21 Distribution of all 22 Time interval and 22 Time interval and 23 Cross tabulation by the results of 24 Estimates of diag 25 Any adverse ever	llection was planned before the index test and reference standard	/	
7 On what basis por (such as symptor) 8 Where and when 9 Whether particip 10b Reference standar 11 Rationale for choose 12a Definition of and of the index test, in suff the index test, in suff of the index test, in suff of the index test, in suff of the index test, in suff and of the index test, in suff the index test, in suff and of the index test, in suff of the inde	(prospective study) or after (retrospective study)	V	5
(such as symptor) 8		\checkmark	5
8 Where and when 9 Whether particip 10 Index test, in suff 10 Reference stands 11 Rationale for cho 12 Definition of and of the index test, 12 Definition of and of the reference 13 Whether clinical to the performer 13 Whether clinical to the assessors of the assessors of the set of the index test, 15 How indetermines 16 How missing data 17 Any analyses of value 18 Intended sample RESULTS Participants 19 Flow of participa 20 Baseline demogration of set of the performer 21 Distribution of all 22 Time interval and 22 Time interval and 24 Estimates of diag 25 Any adverse ever DISCUSSION	tentially eligible participants were identified	¥	
9 Whether particip Test methods 10a Index test, in suff 10b Reference standa 11 Rationale for cho 12a Definition of and of the index test, 12b Definition of and of the reference 13a Whether clinical to the performer 13b Whether clinical to the assessors of the assessors of the index test, 15 How indetermina 16 How missing data 17 Any analyses of will all intended sample RESULTS Participants 19 Flow of participa 20 Baseline demogration of second in the index test, 21b Distribution of all carries are interval and carries tresults 23 Cross tabulation by the results of carries are carried and carried an	ns, results from previous tests, inclusion in registry)	\checkmark	5
10a Index test, in sufficient 10b Reference standa 11 Rationale for choose 12a Definition of and of the index test, 12b Definition of and of the reference 13a Whether clinical to the performer 13b Whether clinical to the assessors of the assessors of the index test 15 How indetermina 16 How missing data 17 Any analyses of value 18 Intended sample	potentially eligible participants were identified (setting, location and d	ates)/	7
10b Reference standar 11 Rationale for cho 12a Definition of and of the index test, 12b Definition of and of the reference of the reference of the performer 13b Whether clinical to the performer 13b Whether clinical to the assessors of the assessors of the reference of the performer of the performer of the assessors of the performer of the perf	ants formed a consecutive, random or convenience series	7	5
11 Rationale for choose 12a Definition of and of the index test, 12b Definition of and of the reference 13a Whether clinical to the performer 13b Whether clinical to the assessors of the assessors of the index test, 15 How indetermine 16 How missing date 17 Any analyses of with 18 Intended sample 18 RESULTS Participants 19 Flow of participa 20 Baseline demogration 21a Distribution of all 22 Time interval and 22 Time interval and 23 Cross tabulation by the results of 24 Estimates of diag 25 Any adverse ever	cient detail to allow replication	Y /	5-6, 8
12a Definition of and of the index test, 12b Definition of and of the reference. 13a Whether clinical to the performer. 13b Whether clinical to the assessors of the assessors of the series of the test of the assessors of the	rd, in sufficient detail to allow replication	./	5, 8
12a Definition of and of the index test, 12b Definition of and of the reference. 13a Whether clinical to the performer. 13b Whether clinical to the assessors of the assessors of the series of the test of the assessors of the	osing the reference standard (if alternatives exist)	Y /	4
of the index test, 12b Definition of and of the reference of the reference of the performer of the performe	rationale for test positivity cut-offs or result categories	······································	_
12b Definition of and of the reference solution of the reference solution of the reference solution in the performer solution to the performer solution to the performer solution to the assessors of the assessors of the assessors of the assessors of the solution in the session of the session solution in the session of the session solution in the session of the sess	distinguishing pre-specified from exploratory	\checkmark	6
of the reference of the	rationale for test positivity cut-offs or result categories		_
to the performer 13b Whether clinical to the assessors of the assessors o	tandard, distinguishing pre-specified from exploratory	\checkmark	4
13b Whether clinical to the assessors of	nformation and reference standard results were available		_
to the assessors of Analysis 14 Methods for esting 15 How indeterminary 16 How missing data 17 Any analyses of wind 18 Intended sample 18 Intended sample 20 Baseline demogration 21 Distribution of second 21 Distribution of all 22 Time interval and 22 Time interval and 23 Cross tabulation by the results of 24 Estimates of diag 25 Any adverse ever 25 DISCUSSION	/readers of the index test	\checkmark	5
Analysis 14 Methods for esting the state of	nformation and index test results were available		E
15 How indeterminate 16 How missing data 17 Any analyses of v 18 Intended sample RESULTS Participants 19 Flow of participa 20 Baseline demograte 21a Distribution of second 21b Distribution of all 22 Time interval and 23 Cross tabulation 24 Estimates of diag 25 Any adverse ever DISCUSSION	f the reference standard	\checkmark	5
16 How missing data 17 Any analyses of v 18 Intended sample RESULTS Participants 19 Flow of participa 20 Baseline demogra 21a Distribution of security and securi	nating or comparing measures of diagnostic accuracy	\	6
17 Any analyses of v 18 Intended sample RESULTS Participants 19 Flow of participa 20 Baseline demogra 21a Distribution of al 21b Distribution of al 22 Time interval and Test results 23 Cross tabulation by the results of al 24 Estimates of diag 25 Any adverse ever	te index test or reference standard results were handled	\	8
18 Intended sample RESULTS Participants 19 Flow of participal 20 Baseline demogra 21a Distribution of se 21b Distribution of al 22 Time interval and 23 Cross tabulation by the results of 3 24 Estimates of diag 25 Any adverse ever	on the index test and reference standard were handled	Y	8
18 Intended sample RESULTS Participants 19 Flow of participal 20 Baseline demogra 21a Distribution of se 21b Distribution of al 22 Time interval and 23 Cross tabulation by the results of 3 24 Estimates of diag 25 Any adverse ever	ariability in diagnostic accuracy, distinguishing pre-specified from explo	ratory	7-8
RESULTS Participants 19 Flow of participa 20 Baseline demogra 21a Distribution of se 21b Distribution of al 22 Time interval and Test results 23 Cross tabulation by the results of 24 Estimates of diag 25 Any adverse ever	size and how it was determined	\\\\	5
20 Baseline demogra 21a Distribution of se 21b Distribution of al 22 Time interval and 23 Cross tabulation by the results of 3 24 Estimates of diag 25 Any adverse ever		. .	
20 Baseline demogra 21a Distribution of se 21b Distribution of al 22 Time interval and 23 Cross tabulation by the results of 3 24 Estimates of diag 25 Any adverse ever	nts, using a diagram		Fig.1 + Fig.
21a Distribution of section 21b Distribution of all 22 Time interval and 23 Cross tabulation by the results of 24 Estimates of diag 25 Any adverse ever DISCUSSION	phic and clinical characteristics of participants	\ <u>\</u>	6
21b Distribution of all 22 Time interval and 23 Cross tabulation by the results of 24 Estimates of diag 25 Any adverse ever DISCUSSION	verity of disease in those with the target condition	Y /	6
22 Time interval and	ernative diagnoses in those without the target condition		6
Test results 23 Cross tabulation by the results of 24 Estimates of diag 25 Any adverse ever DISCUSSION	any clinical interventions between index test and reference standard	~	6
by the results of 24 Estimates of diag 25 Any adverse ever DISCUSSION	of the index test results (or their distribution)	· V .	
25 Any adverse ever DISCUSSION	he reference standard	\checkmark	8
25 Any adverse ever DISCUSSION	nostic accuracy and their precision (such as 95% confidence intervals)	./	8
DISCUSSION	ts from performing the index test or the reference standard	<u>-</u>	8
		···V	
: Zo : Study limitations	including sources of potential bias, statistical uncertainty, and generali	sability	9
	ractice, including the intended use and clinical role of the index test	. /	9
OTHER 27 Implications for p			
	per and name of registry	<i>I</i>	11
	idy protocol can be accessed	~ / ,	11
	g and other support; role of funders	✓,	10-11



STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A diagnostic accuracy study evaluates the ability of one or more medical tests to correctly classify study participants as having a target condition. This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

