Section & Topic		No	Item
TITLE OR ABSTRACT			
	OK	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)
ABSTRACT			
	OK	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)
INTRODUCTION			
	OK	3	Scientific and clinical background, including the intended use and clinical role of the index test
	OK	4	Study objectives and hypotheses
METHODS			
Study design	OK	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)
Participants	OK	6	Eligibility criteria
	OK	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)
	OK	8	Where and when potentially eligible participants were identified (setting, location and dates)
	OK	9	Whether participants formed a consecutive, random or convenience series
Test methods	OK	10a	Index test, in sufficient detail to allow replication
		10b	Reference standard, in sufficient detail to allow replication
		11	Rationale for choosing the reference standard (if alternatives exist)
		12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory
		12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory
		13a	Whether clinical information and reference standard results were available to the performers/readers of the index test
		13b	Whether clinical information and index test results were available to the assessors of the reference standard
Analysis	OK	14	Methods for estimating or comparing measures of diagnostic accuracy
	OK	15	How indeterminate index test or reference standard results were handled
	OK	16	How missing data on the index test and reference standard were handled
	OK	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory
	OK	18	Intended sample size and how it was determined
RESULTS			
Participants	OK	19	Flow of participants, using a diagram
	OK	20	Baseline demographic and clinical characteristics of participants
	OK	21a	Distribution of severity of disease in those with the target condition
		21b	Distribution of alternative diagnoses in those without the target condition
		22	Time interval and any clinical interventions between index test and reference standard
Test results	OK	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard
	OK	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)
		25	Any adverse events from performing the index test or the reference standard
DISCUSSION			
	OK	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability
	OK	27	Implications for practice, including the intended use and clinical role of the index test
OTHER INFORMATION			
		28	Registration number and name of registry
		29	Where the full study protocol can be accessed
	OK	30	Sources of funding and other support; role of funders

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.

Figure S1 - TRIPOD Checklist: Prediction Model Validation

Section/Topic	Item		Checklist Item	
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	ok
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	ok
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	ok
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	ok
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	ok
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	ok
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	ok
	5b	D;V	Describe eligibility criteria for participants.	ok
	5c	D;V	Give details of treatments received, if relevant.	
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	ok
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	ok
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	ok
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	ok
Sample size	8	D;V	Explain how the study size was arrived at.	ok
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	ok
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	
	10c	V	For validation, describe how the predictions were calculated.	ok
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	ok
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	ok
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	ok
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	ok
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	ok
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	
	15b	D	Explain how to the use the prediction model.	
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	ok
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	ok
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	ok
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	ok
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	ok
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	ok