Supplementary Materials S1: SCOLIOSIS SELF-DETECTION FACT SHEET

Here is an explanation of how to detect scoliosis.

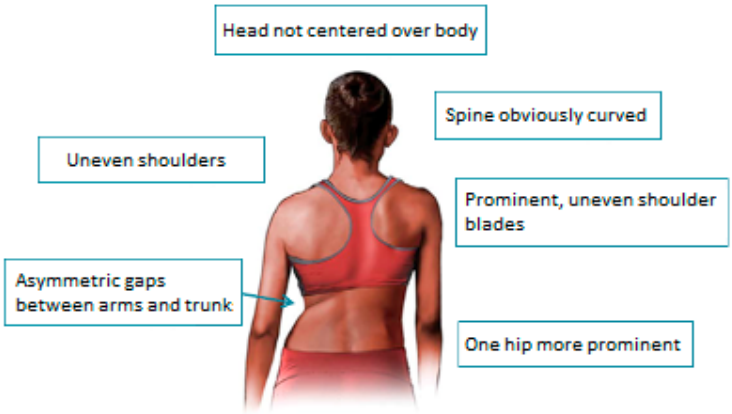
What is scoliosis?

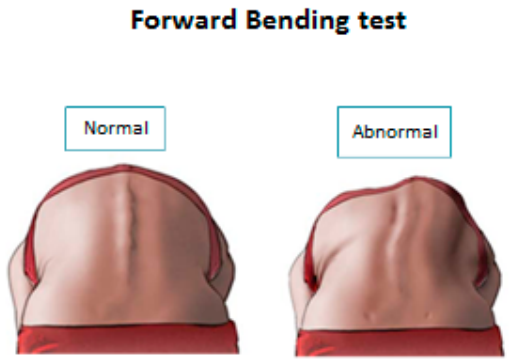
Scoliosis is a lateral or sideways curve of the spine. The spine also rotates on its long axis as it curves. It usually develops during early adolescence (between 10 and 18 years) when growth is most rapid.

What are the outward signs of scoliosis?

Scoliosis can be detected by looking closely at the back of your child. Due to the curvature of the spine and the rotation of the spinal vertebrae, various external characteristics arise:

-Head not centered over the body
-Uneven shoulders
-Prominent, uneven shoulder blades
-Asymmetric gaps between arms and trunk
-Spine obviously curved
-One hip more prominent
-Bending test reveals an uneven portion of the back





Not all features are always clearly and visibly present in all patients, nor do all features have to be present together.

Supplementary Materials S2: Characteristics of included patients in the survey

CHARACTERISTICS OF AIS PATIENTS INCLUDED IN THE SURVEY

Survey 1

Patient Number	Category	Sex	Age (y)	Upper level	Lower level	Cobb angle
1	No scoliosis	M	12	x	x	x
2	T10	M	14	T1	T5	18°
3	T20	M	14	T6	t12	28°
4	L20	F	11	L1	L4	24°
5	T10	F	16	T6	T10	19°
6	T30	M	16	T8	T12	35°
7	No scoliosis	M	15	x	x	X
8	No scoliosis	M	14	x	x	X
9	L30	F	15	T11	L4	30°
10	No scoliosis	M	15	x	x	X
11	T20	F	13	T6	T11	22°
12	No scoliosis	F	16	x	x	X
13	No scoliosis	F	9	x	x	X
14	T30	F	14	T6	T11	31°

Survey 2

Patient Number	Category	Sex	Age (y)	Upper level	Lower level	Cobb angle
1	T20	F	14	T6	T11	24°
2	L30	F	14	T12	L3	35°
3	No scoliosis	F	16	x	x	X
4	L20	F	16	T11	L4	23°
5	No scoliosis	M	14	x	x	X
6	No scoliosis	F	9	x	x	X
7	No scoliosis	M	15	x	x	X
8	T10	M	15	T6	L1	19°
9	T30	F	13	T5	L1	33°
10	T30	F	14	T6	T11	37°
11	No scoliosis	M	12	x	x	X
12	T20	F	14	T6	T10	25°
13	No scoliosis	M	15	x	x	X
14	T10	F	15	T7	T12	16°

Abbreviations: T10, Thoracic 10°-20°; T20, Thoracic 20°-30°; T30, Thoracic 30°-40°; L20, Lumbar 20°-30°; L30, Lumbar 30°-40°; M, Male; F, Female

Number of patients included per category

Category	Amount of patients
No scoliosis	6
T10	2
T20	2
T30	2
L20	1
L30	1

Supplementary Materials S3: STARD Checklist

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	3
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	3
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5-6
Participants	**6**	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random, or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6
	15	How indeterminate index test or reference standard results were handled	6
	16	How missing data on the index test and reference standard were handled	There are no missing data, because all the assessors completed the survey.
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6
	18	Intended sample size and how it was determined	6
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Not available and not relevant for this study. Instead we provided important characteristics of both the assessors and the assessed patients.
	20	Baseline demographic and clinical characteristics of participants	6, 12
	21a	Distribution of severity of disease in those with the target condition	17
	21b	Distribution of alternative diagnoses in those without the target condition	17
	22	Time interval and any clinical interventions between index test and reference standard	No time interval.

Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	13
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	13
	25	Any adverse events from performing the index test or the reference standard	No adverse events.
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalizability	8-9
	27	Implications for practice, including the intended use and clinical role of the index test	9
OTHER INFORMATION			
	28	Registration number and name of registry	1
	29	Where the full study protocol can be accessed	1
	30	Sources of funding and other support; role of funders	1

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 the **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, the categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple of 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.