SPINE Volume 23, Number 7, pp 796–803

A Study of the Diagnostic Accuracy and Reliability of the Scoliometer and Adam's Forward Bend Test

Pierre Côté, DC, MSc, FCCSC,* Brad G. Kreitz, DC, FCCSC,‡
J. David Cassidy, DC, PhD, FCCSC,* Anne K. Dzus, MD, FRCSC,†
and Johanne Martel, BSc, DC‡

Study Design. Study of the diagnostic accuracy and interexaminer reliability of scoliosis diagnostic tests.

Objectives. To estimate the sensitivity, specificity, and predictive value of the Scoliometer (National Scoliosis Foundation, Watertown, MA) and Adam's forward bend test in diagnosing scoliosis, and to determine the interexaminer reliability of the Scoliometer and Adam's forward bend test.

Summary of Background Data. Exposure to diagnostic radiation in patients with adolescent idiopathic scoliosis may result in a small but significant increase in cancer rates. The full-spine radiographic examination remains the standard procedure for the assessment of scoliosis. There is a need for a valid and reliable noninvasive test to assess scoliosis.

Methods. Two examiners independently assessed 105 patients presenting to a scoliosis clinic for trunk asymmetry with Adam's forward bend test and axial trunk rotation with the Scoliometer. The Cobb method served as the gold standard.

Results. The interexaminer agreement for the Scoliometer is excellent in the thoracic spine and substantial in the lumbar spine. The interexaminer measurement error shows poor precision for thoracic and lumbar Scoliometer measurements. The interexaminer agreement for Adam's forward bend test is substantial in the thoracic spine and poor in the lumbar spine. Adam's forward bend test is more sensitive than the Scoliometer in detecting thoracic curves measuring 20° or more by the Cobb method. Receiver operating characteristic curve analysis suggests that the use of the Scoliometer marginally improves the ability of diagnosing a scoliosis in the thoracic spine.

Conclusions. The Scoliometer and Adam's forward bend tests have adequate interexaminer reliability for the assessment of thoracic curves. The Scoliometer has better interexaminer agreement in the lumbar spine. However, the Scoliometer has a high level of interexaminer measurement error that limits its use as an outcome instrument. Because Adam's forward bend test is more sensitive than the Scoliometer, the authors believe that it remains the best noninvasive clinical test to evaluate scoliosis. [Key words: Adam's forward bend test, diagnosis, idiopathic scoliosis, reliability, scoliometer, validity] **Spine 1998;23:796–803**

Adolescent idiopathic scoliosis predominantly affects girls during the peak of their physiologic growth, when their maturing tissues are most vulnerable to radiation exposure.[19,34] Because the gold standard procedure to diagnose and follow the progression of scoliosis is the Cobb method of radiographic measurement, these young patients can be exposed to a significant amount of diagnostic radiation. Several studies have demonstrated that the use of multiple full-spine radiographs results in a small, but significant increase in the incidence of cancer in the adolescent idiopathic scoliosis population.[11,21,23,30] Levy et al[23] estimated that the excess lifetime risk for development of cancer after an average of 12 scoliosis radiographs is 1% to 2% (12–25/1000) in women. Although this represents a potential public health issue, there are no valid, reliable, noninvasive alternative methods to assess and monitor the progression of scoliosis.

Many attempts have been made to validate noninvasive tests used to assess scoliosis.[1,6,28,31] Bunnel[6] postulated that the Scoliometer (National Scoliosis Foundation, Watertown, MA), an inclinometer used to measure axial trunk rotation during forward bending, could document curve progression and consequently be a practical way of decreasing exposure to x-rays. He described the Scoliometer as a valid screening instrument to evaluate curve magnitude. However, a careful look at his data reveals that the sensitivity and specificity of the Scoliometer are only 23% and 48%, respectively. Another validity study reported that the sensitivity and specificity of the Scoliometer vary with the selected cut-off points. Amendt et al[1] suggest that with axial trunk rotation levels of 5° or more, the Scoliometer is 98% sensitive and 29% specific in detecting curves measuring 20° or more with the Cobb method. However, the sensitivity decreases to 51% and the specificity increases to 96% when the cut-off point is 10°.

From the *Institute for Health and Outcomes Research, Department of Physical Medicine and Rehabilitation, and the †Department of Surgery, University of Saskatchewan, Saskatoon; and ‡Private Chiropractic Practice, Montreal, Quebec, Canada.
Supported by the Chiropractors' Association of Saskatchewan.
Acknowledgment date: August 13, 1996.
Acceptance date: July 29, 1997.
Device status category: 9.

Two studies suggest that the Scoliometer has excellent interexaminer agreement in assessing axial trunk rotation in the thoracic and lumbar spine.[1,28] However, the presence of methodologic and analytic flaws restricts the generalizability of these studies. Murrell et al,[28] in a study conducted on a small sample of 22 female adolescents, failed to report the characteristics of their subjects, thereby limiting the external validity of their findings. In another study, Amendt et al[1] investigated 65 scoliotic subjects and reported high interexaminer reliability. However, their conclusions may be inaccurate because they used an inappropriate statistic to measure interexaminer reliability. Pearson's "r" correlation coefficient is a measure of association and tends to overestimate agreement.[5] The intraclass correlation coefficient (ICC) is the proper method to assess agreement in interval data.[3]

A crucial, but often neglected component of reliability analysis is the quantification of the interexaminer measurement error (IME), or precision. The IME is a measure of disagreement between examiners that can be expressed as a standard deviation or confidence interval. In the only study commenting on the interexaminer error associated with Scoliometer readings, Murrell et al[28] report a standard deviation of 1.9° for the thoracic spine and 2.3° for the lumbar spine. Assuming that the measurement errors are normally distributed, the current authors computed that it corresponds to 95% confidence intervals of 0° to 5.3° for the thoracic spine and 0° to 6.4° for the lumbar spine. These levels of IME are extremely high, considering that 5° of axial trunk rotation is considered necessary to detect a Cobb angle measurement of 20° or more.[6]

Although controversial evidence exists regarding the reliability and validity of the Scoliometer, it has been recommended as an adequate clinical tool and outcome measure for scoliosis research.[26,33] Furthermore, it has been used in a large, prospective trial of the effectiveness of bracing for adolescent idiopathic scoliosis.[29] In this study, the authors investigate the reliability and diagnostic accuracy of the Scoliometer and Adam's forward bend test in the evaluation of patients with adolescent idiopathic or congenital scoliosis presenting to a tertiary center.

■ Methods

Sample Size Estimate. A sample size estimate was computed for the reliability component of the study.[32] To obtain an ICC of 0.7 with a 95% confidence interval of 0.2 at an alpha level of 0.05, 99 subjects were required.

Definition of Sample. All consecutive patients referred to a university hospital pediatric scoliosis clinic for evaluation of adolescent idiopathic or congenital scoliosis were eligible for the study. Although patients with congenital scoliosis are more likely to have true lateral curves with little rotation, they were included to test the Scoliometer and Adam's forward bend test on a wide range of curve pattern and magnitude. The authors also included patients who had undergone scoliosis surgery and those who wore a brace at the time of the investigation because of their need to be radiographed throughout their treatment. Patients were excluded who presented with scoliosis of other known etiologies, as were those with idiopathic or congenital curves who were unable to perform the forward bend test. Further, patients who were radiographed in a brace, or who had a radiograph that would not allow proper measurement of the Cobb angles, also were excluded.

Experimental Procedure. Two residents performed the Scoliometer measurements and Adam's forward bend test. To standardize their method of assessment, the examiners participated in a training session that took place before the outset of the study. The Scoliometer measurements were performed independently and by convenience, with each examiner being blinded to the other's measurements. To reduce measurement error, the examiners used the same instrument to assess all patients.

The Scoliometer is a fluid-filled inclinometer in which an enclosed ball shows the angle of trunk rotation on a scale of 1° increments that ranges from 0° to 30°. To improve the conformity of measurements, the angle of trunk rotation recorded is the one corresponding to the highest value entirely crossed by the enclosed ball. The examiners placed the Scoliometer on the thoracic or lumbar spine while the patient performed Adam's forward bend test. They instructed the patients to place their feet shoulder-width apart, extend their arms with the hands placed together, and bend forward slowly. If this position revealed trunk asymmetry, they recorded Adam's forward bend test as positive. Depending on the location of the thoracic rotational prominence or lumbar flank prominence, the examiners asked the patients to flex forward until the hump approximated the horizontal plane. The examiners placed the Scoliometer at the level of the most prominent aspect of the hump, with the "0" mark centered over the spine. The estimated apical segment, degree of trunk rotation, and the side of the hump were recorded. Scoliometer measurements were obtained from the thoracic and lumbar regions on all patients.

Standing full-spine posteroanterior and lateral scoliosis radiographs were obtained for each new patient referred to the clinic. Patients returning to the clinic for follow-up were radiographed in the posteroanterior position. All patients had their Cobb angle measured by the same experienced pediatric orthopedic surgeon. The authors selected the Cobb method as their gold standard because of its good reliability in measuring scoliosis.[14] To minimize instrument measurement error, the pediatric orthopedic surgeon measured the spinal curvatures with the same scoliosis ruler throughout the study. For the purposes of this study, patients with more than two curves had the uppermost and lowest curves measured. All data were entered into a computer database by an independent study member.

Statistical Analysis. The authors estimated the interexaminer reliability for Scoliometer readings and location of apical vertebrae with the ICC and its 95% confidence interval (CI).[2,3,15,16,27] A two-way, random-effects analysis of variance model using the examiners as a random sample was used to compute the ICC. Furthermore, the magnitude of disagreement between the examiners was determined for the Scoliometer measurements by computing the IME and relative IME.[15] For

the nominal data, the Cohen's kappa (k) coefficient and its 95% CI were used to assess consistency between the two examiners.[12,15]

The authors calculated the crude sensitivity, specificity, and predictive value estimates of the Scoliometer and Adam's forward bend test from the uppermost and lowest curves individually using contingency table analysis. The precision of the estimates is reported as 95% CI.[4] For the analysis, the Scoliometer measurements were dichotomized by setting a positive test at 5° or more of trunk rotation.[6] A scoliosis was considered present if and only if the two examiners agreed that the Scoliometer or Adam's forward bend test were positive. In other words, the two examiners had to detect a thoracic rotational prominence or lumbar flank prominence, or obtain axial trunk rotation measurements above or equal to 5° for a curve to be categorized as present. Similarly, the Cobb angle measurements were dichotomized with a cut-off point set at 20°.

The diagnostic accuracy of a clinical test depends on the characteristics of the study sample.[20,25] For example, the ability of the Scoliometer to measure axial trunk rotation accurately may vary with the anthropometric differences between male and female patients. To measure the impact of such differences, the authors determined the strata-specific sensitivity and specificity estimates in their study for age group, gender, and history of surgery using logistic regression analysis.[8] This approach is used to obtain smoothed estimates when a small number of cases are present in specific strata. In the regression model, the dependent variable was defined as the dichotomized result of the screening test: Scoliometer or Adam's forward bend test. The gold standard (Cobb angle) was used as a binary independent variable. The authors built different models to assess the independent effect of age group, gender, and history of surgery. The two age groups were dichotomized with a cut-off point set at younger than 14 years of age, or 14 years of age or older. This cut-off point was selected for its physiologic value, because it is known that curves measuring less than 19° are unlikely to progress in adolescents older than 13–15 years of age.[24] The β coefficients obtained from the regression model were used to estimate the strata-specific sensitivity and specificity rates. The 95% confidence limits for each covariate level were determined from the estimated overall variance of each strata-specific logit.

Finally, receiver operating characteristic (ROC) curve analysis was used to 1) examine the trade-off between sensitivity and specificity when different Scoliometer cut-off points were used; 2) determine the probability that a patient with a scoliosis measuring more than 20° with the Cobb method would be correctly identified by the Scoliometer; and 3) compare the performance of the Scoliometer in the thoracic and lumbar spine in detecting curves measuring 20° or more with the Cobb method. The ROC curve analysis was conducted according to the trapezoid method described by Hanley and McNeil.[17,18]

■ Results

One hundred sixty-three patients presented to the scoliosis clinic between February 22 and June 21, 1994. One hundred five patients met the inclusion criteria. Fifty-eight patients were excluded: 24 had neuromuscular or pathologic curves, 15 could not adequately bend forward, 8 had infantile or juvenile idiopathic scoliosis, 6 were radiographed in their brace, 1 lacked radiographic landmarks that could be accurately established, and 4 were excluded for other reasons.

Table 1. Descriptive Statistics for the Scoliometer for Examiners 1 and 2 (N = 105)

	Mean (°)	SD (°)	Range (°)
Thoracic spine			
Examiner 1	6.3	5.5	0–28
Examiner 2	6.1	5.6	0–30
Lumbar spine			
Examiner 1	4.3	3.6	0–14
Examiner 2	5.5	4.5	0–20

There are 87 girls and 18 boys in the final study sample. The mean age of the study sample is 15.5 years (SD = 4.8). Ten patients had undergone corrective surgery for their scoliosis, and 16 were treated with a brace. Two patients had congenital curves. The mean magnitude of the curves measured by the Cobb method was 22.6° (SD = 16.3°) for the thoracic curves and 19.5° (SD = 13.9°) for the lumbar curves. Seventy percent of the thoracic curves had their apex between T8 and T12, and 77% of the lumbar curves had their apex between L1 and L3. The number of curves on radiographs varied across the sample: 2 patients had no deformity, 16 had one curve, 73 had two curves, and 14 had three curves.

The descriptive statistics for the Scoliometer measurements for Examiners 1 and 2 are presented in Table 1. In the thoracic region, the ICC (95% CI) for the Scoliometer interexaminer agreement is 0.91 (0.86–0.94), which represents excellent agreement beyond chance.[12,22] The IME is 4.9°, indicating that the authors are 95% confident that the Scoliometer measurements did not vary by more than 4.9°. The relative IME is 79%, which indicates that almost 80% of the average measurements can be accounted for by the interexaminer error. For the lumbar curves, the ICC (95% CI) is 0.74 (0.58–0.84), which corresponds to substantial agreement beyond chance. The IME is 6.6° and the relative IME is 134%. In the lumbar spine, more than 100% of the average measurement can be accounted for by the IME alone.

The Cohen's kappa coefficients (k) and 95% CI for the interexaminer agreement of the dichotomized Scoliometer results show substantial agreement in the thoracic spine (k = 0.70 [0.56–0.83]) and fair agreement in the lumbar spine (k = 0.48 [0.32–0.62]).

Table 2 presents the frequency statistics for Adam's forward bend test for both examiners. Adam's forward

Table 2. Positive Adam's Test for Examiners 1 and 2 (N = 105)

	Thoracic Spine (%)	Lumbar Spine (%)
Examiner 1	72.4	52.4
Examiner 2	74.3	81.0

Table 3. Crude Sensitivity, Specificity, and Predictive Value Estimates (95% CI) of the Scoliometer and Adam's Forward Bend Test for Thoracic Curves Using the Cobb Angle as a Gold Standard

Estimate	Scoliometer [% (CI)]	Adam's Test [% (CI)]
Sensitivity	71 (59, 84)	92 (85, 100)
Specificity	83 (73, 93)	60 (47, 74)
Positive predictive value	80 (69, 92)	70 (59, 80)
Negative predictive value	75 (64, 86)	80 (79, 99)

CI = confidence interval.

Table 4. Crude Sensitivity, Specificity, and Predictive Value Estimates (95% CI) of the Scoliometer and Adam's Forward Bend Test for Lumbar Curves Using the Cobb Angle as a Gold Standard

Estimate	Scoliometer [% (CI)]	Adam's Test [% (CI)]
Sensitivity	51 (47, 55)	73 (60, 86)
Specificity	83 (74, 93)	68 (57, 80)
Positive predictive value	70 (54, 95)	64 (49, 79)
Negative predictive value	69 (59, 80)	77 (66, 89)

CI = confidence interval.

bend test interexaminer reliability coefficient (k) is 0.61 (95% CI, 0.44–0.78) for the detection of a thoracic rotational prominence, and 0.29 (95% CI, 0.14–0.44) for the lumbar flank prominence. Therefore, substantial agreement beyond chance is present for the thoracic hump, but poor agreement is obtained for the lumbar flank prominence.

Poor interexaminer agreement is found for clinically determining the location of the apical vertebrae, with an ICC of 0.25 (95% CI, 0.05–0.42) for the thoracic curve apex and 0.28 (95% CI, 0.09–0.45) for the lumbar apex. Therefore, it was difficult for the examiners to agree clinically on the location of the curve apex.

Table 3 presents the crude sensitivity, specificity, and predictive value estimates for the Scoliometer and Adam's forward bend test in the evaluation of the thoracic curves. Overall, the Scoliometer has a crude sensitivity of 71% (95% CI, 59–84) and specificity of 83% (95% CI, 73–93) for detecting thoracic curves measuring 20° or more with the Cobb method. On the other hand, Adam's forward bend test is 92% (95% CI, 85–100) sensitive and 60% (95% CI, 47–74) specific in assessing thoracic curves. As evidenced by the lack of overlap in the 95% CIs, Adam's forward bend test is more sensitive than the Scoliometer. However, the Scoliometer is more specific, yielding a lower proportion of false-positive results. Consequently, the predictive value positive of the Scoliometer is slightly superior. In other words, the probability of accurately identifying a thoracic scoliosis was 80% for the Scoliometer compared with 70% for Adam's forward bend test.

The Scoliometer is 51% (95% CI, 47–55) sensitive and 83% (95% CI, 74–93) specific in detecting lumbar curves measuring 20° or more by the Cobb method (Table 4). In contrast, Adam's forward bend test is 73% (95% CI, 60–86) sensitive and 68% (95% CI, 57–80) specific. The findings obtained for the thoracic curves are reproduced in the lumbar spine. Adam's forward bend test is more sensitive, but it also yields a higher false-positive rate, which lowers its predictive value positive.

Table 5 gives the age-, gender-, and history-of-surgery–specific sensitivity and specificity estimates for the thoracic curves. Interestingly, the sensitivity and specificity of both methods do not vary with age and gender. However, the Scoliometer is more sensitive (85%; 95% CI, 66–94) in evaluating thoracic curves of subjects with previous surgery than subjects without surgery (68%; 95% CI, 40–88). This difference may reflect the influence of the severity of the curve on the estimates or even suggest the presence of examiner bias, introduced by the observation of the surgical scar. The results obtained from the crude analysis replicate themselves in the stratified analysis. Overall, the Scoliometer is more specific than Adam's forward bend test in evaluating thoracic curves. This difference is particularly evident in patients who have not required surgery.

In the lumbar spine, stratification by age group, gender, or history of surgery did not show any confounding effect of these variables on the crude estimates (Table 6). The stability of these estimates suggests that the accuracy of the two tests does not vary with the age group and gender characteristics of the patients. Again, Adam's forward bend test is more sensitive but less specific than the Scoliometer.

Figure 1 shows the ROC curve describing the diagnostic accuracy of the Scoliometer in the thoracic spine. The ROC curve provides a graphic representation of the trade-off that exists between sensitivity and specificity. For example, in the upper right corner of Figure 1 is a point showing the sensitivity and specificity of the Scoliometer when using a cut-off point of 1°. Because 100%

Table 5. Strata-Specific Sensitivity and Specificity Estimates (95% CI) of the Scoliometer and Adam's Forward Bend Test for Thoracic Curves Using the Cobb Angle as a Gold Standard

	Scoliometer [% (CI)]		Adam's Test [% (CI)]	
	Sensitivity	Specificity	Sensitivity	Specificity
Age group				
<14 yr	67 (35, 89)	86 (69, 94)	89 (65, 97)	70 (49, 85)
≥14 yr	73 (41, 91)	82 (62, 92)	94 (78, 98)	55 (47, 74)
Gender				
Male	71 (44, 89)	83 (72, 90)	90 (72, 96)	69 (66, 73)
Female	71 (43, 89)	83 (70, 91)	93 (78, 98)	59 (45, 72)
Surgery				
Yes	85 (66, 94)	65 (53, 76)	95 (86, 96)	59 (45, 72)
No	68 (40, 88)	83 (71, 91)	92 (75, 98)	50 (36, 64)

CI = confidence interval.

Table 6. Strata-Specific Sensitivity and Specificity Estimates (95% CI) of the Scoliometer and Adam's Forward Bend Test for Thoracic Curves Using the Cobb Angle as a Gold Standard

	Scoliometer [% (CI)]		Adam's Test [% (CI)]	
	Sensitivity	Specificity	Sensitivity	Specificity
Age group				
<14 yr	51 (23, 77)	83 (68, 92)	76 (49, 91)	66 (47, 81)
≥14 yr	51 (23, 79)	83 (68, 92)	72 (45, 89)	70 (51, 83)
Gender				
Male	59 (29, 83)	80 (63, 90)	66 (41, 85)	75 (62, 85)
Female	50 (24, 76)	84 (81, 92)	74 (51, 89)	67 (63, 70)
Surgery				
Yes	43 (32, 55)	88 (85, 91)	64 (45, 80)	79 (70, 85)
No	53 (25, 80)	83 (72, 91)	76 (52, 80)	68 (55, 79)

CI = confidence interval.

of the patients had at least 1° of rotation, the proportion of true-positive results (sensitivity) is 100% when using this cut-off point. However, the corresponding proportion of false-positive results (1-specificity) for this cut-off point is high at 85%. As the cut-off point is increased, the curve moves to the left. This transition produces a sharp increase in specificity and a decrease in sensitivity. In fact, at a cut-off point of 12°, the specificity reaches 100% and the sensitivity declines to 23%. In other words, 77% of the patients with a scoliosis were misclassified as false negatives, but there are no false-positive results. The ROC curve for the thoracic spine lies well above the diagonal. The area under the curve (AUC) for the thoracic spine is 0.719 (SE = 0.032). As determined by the Wilcoxon rank-sum test,[15] this is significantly

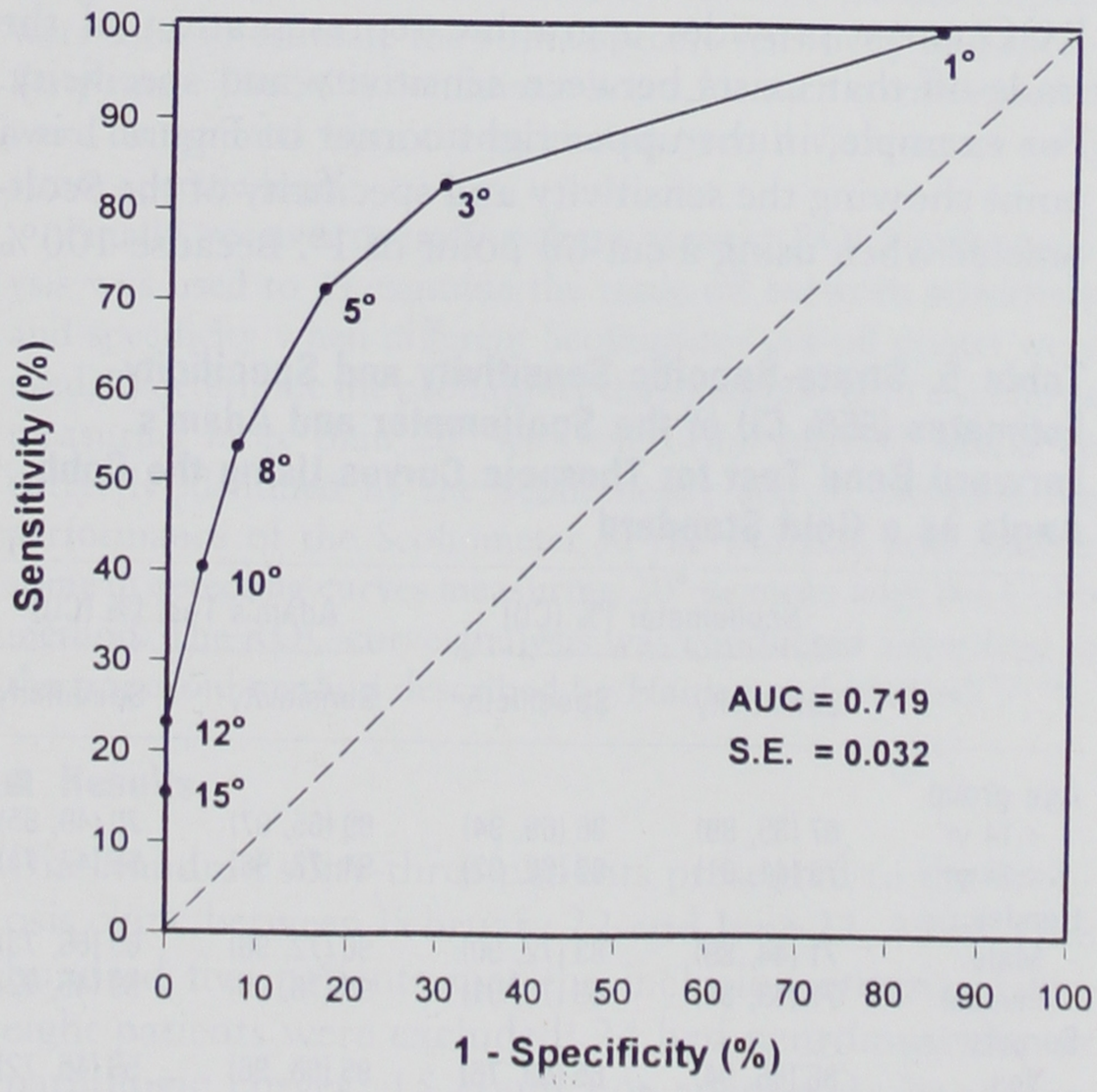

Figure 1. Receiver operating characteristic curve for thoracic scoliosis measurements with the Scoliometer.

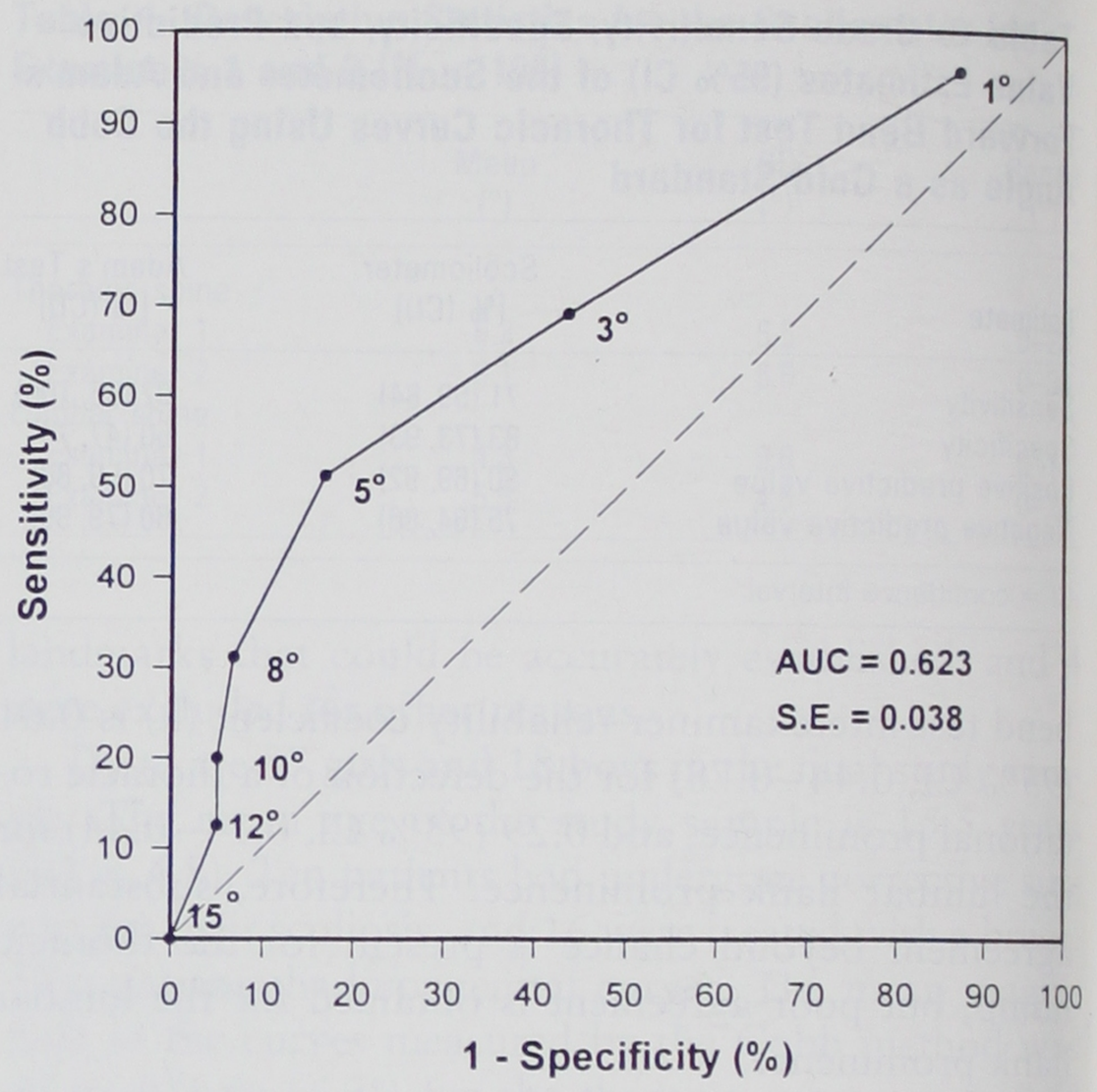

Figure 2. Receiver operating characteristic curve for lumbar scoliosis measurements with the Scoliometer.

($P = 0.018$) greater than an AUC of 0.5, which corresponds to the line of no information (diagonal).

In the lumbar spine, the ROC curve approximates the diagonal more closely. An abrupt decrease in sensitivity is associated with a sharp increase in specificity as the cut-off point is increased from 1° to 15° (Figure 2). The AUC is 0.623 (SE = 0.038), which is not significantly ($P = 0.17$) greater than the line of no information.

■ Discussion

The results of the current study suggest that the reliability of the Scoliometer is questionable. Although the instrument has good interexaminer agreement, its high level of interexaminer error restricts its usefulness as a diagnostic screening tool. The error associated with Scoliometer measurements is at least 4.9°, regardless of the region examined. Clinically, this corresponds to roughly 100% of the measurement necessary to define the presence of scoliosis according to the Bunnel[6] criteria for specialist referral. Consequently, such a high level of error makes the significance of any measurement trivial. Furthermore, our results corroborate those of Murrell et al,[28] who reported a similar level of interexaminer error.

The high level of interexaminer error jeopardizes the validity of the Scoliometer as an outcome measure in follow-up studies. The authors believe that follow-up measurements of less than 5° in the thoracic spine or 6.5° in the lumbar spine could not be conclusive of curve progression. This strongly suggests that the Scoliometer would not be responsive to clinically important changes in curve magnitude during the follow-up period.

The high level of interexaminer error associated with Scoliometer measurements can be attributed to several

factors. First, the precision of the instrument itself is a potential source of measurement error. The mercury ball used to read the angle of axial trunk rotation is relatively large compared with the increments on the instrument's scale. Second, the apex of the curve, which in general corresponds to the point of maximal trunk rotation, was not identified in a consistent manner by the examiners. This likely influenced the placement of the instrument, which in turn introduced variation in the Scoliometer readings. Third, the current results suggest that it may be difficult for two examiners to perform Adam's forward bend test in a highly consistent manner. The reliability coefficients obtained for Adam's forward bend test (k = 0.61 in the thoracic spine and 0.29 in the lumbar spine) suggest that this method may be liable to measurement bias. It may be argued that the amplitude of forward flexion achieved by the same patient on repeated measurements (Examiner 1 and Examiner 2) may have been different, or that the examiners used a different reference point to report a thoracic rotational prominence. This would influence the height and size of the thoracic and lumbar humps and affect the placement of the instrument. The authors believe that the Scoliometer's level of IME would decrease if the examiners would perform Adam's forward bend test more consistently and if they located the apical vertebra more reliably.

The purpose of developing a noninvasive test to diagnose scoliosis is primarily to reduce the number of required radiographs. Because the consequences of misdiagnosing a scoliosis that measures 20° or more can be devastating for the future health of a skeletally immature patient, a noninvasive test must be very sensitive. In other words, a valid test must be efficient at detecting curves with Cobb angles of 20° or more and yield few false-negative results. Ideally, this test could serve as the first diagnostic procedure to be performed in a serial testing approach.[13] In other words, if this noninvasive test is positive, a full-spine radiograph would be obtained. This would ultimately reduce the number of patients subjected to radiation. Therefore, the authors believe that the Scoliometer should be at least 90% sensitive in detecting curves measuring 20° or more. The current results demonstrate that the Scoliometer does not fulfill this important criterion. In fact, in the thoracic spine, its crude level of sensitivity (71%) is significantly lower than that of Adam's forward bend test (92%). This suggests that Adam's forward bend test remains a better first-line procedure in the assessment of thoracic scoliosis.

The current results differ from those reported by Amendt et al.[1] Using a cut-off point of 5° of axial trunk rotation, they found the Scoliometer to be at least 94% sensitive in detecting curves of 20° or more. However, a review of the characteristics of their sample suggests that their patients may have had curves of greater magnitude than those in the current study. Because severity of disease influences the sensitivity and specificity of a test, their results probably reflect the attributes of their sample. A study of a diagnostic test must include a spectrum of cases ranging from no spinal curvature to severe scoliosis. With 50% and 43%, respectively, of the current patients having curves measuring 20° or more in the thoracic and lumbar spine, the authors are confident that their sample included a wide range of scoliosis severity.

The sensitivity and specificity estimates do not vary with age and gender. This finding implies that both the Scoliometer and Adam's forward bend test are stable procedures in the current sample. The larger proportion of false-positive cases for Adam's forward bend test in the nonsurgical subgroup likely reflects the impact of disease severity on the specificity estimate. Because patients who have not undergone surgery have less severe disease, it may make the clinical appreciation of curve severity more difficult and increase the measurement and random error. Furthermore, the presence of a surgical scar may have biased the decision of the examiners to detect scoliosis in those who underwent surgery.

The ROC curve of a test that possesses good diagnostic accuracy tends to crowd toward the left upper corner of the graph and away from the diagonal. This produces a large AUC that indicates that the test discriminates well between the "presence" and "absence" of the disease. The AUC provides a global estimate of test accuracy, regardless of the cut-off point selected. The authors found that the Scoliometer marginally improves the accuracy of diagnosing scoliosis measuring 20° or more by the Cobb method. Overall, increasing the cut-off point resulted in an unacceptable decrease in sensitivity. Although the Scoliometer significantly increased the ability to diagnose a thoracic scoliosis by 22% over chance alone, it is associated with a large number of false-negative results.

The most important factor that limits the generalizability of the conclusions reached in previous validity studies of diagnostic procedures in tertiary centers is the presence of a sample selection bias. This bias, also known as pretest bias, violates the assumption of independent observation of Bayesian statistics, on which the concept of sensitivity and specificity were developed.[9,10] In the current study, patients referred to the scoliosis clinic were previously diagnosed with a spinal curvature by another physician. These preselected patients were therefore more likely to obtain a definitive diagnosis of scoliosis measuring 20° or more with subsequent testing. The results of the Scoliometer and Adam's forward bend test are therefore not independent from a previous clinical examination. This may have violated the assumption of independence of observations and may restrict the generalizability of the results to tertiary clinics only. Consequently, the conclusions of this study should not be generalized to screening programs or general medical practice.

■ Conclusion

The Scoliometer possesses adequate interexaminer reliability. However, its poor precision and inadequate diagnostic accuracy limit its clinical usefulness as a outcome instrument. The objective quantification of trunk asymmetry with the Scoliometer adds little benefit to the diagnosis of scoliosis in a tertiary clinic. In addition, the instrument is unlikely to be responsive in monitoring curve progression or in reducing the number of diagnostic radiographs. Based on the current results, the authors recommend that Adam's forward bend test and the full-spine radiograph remain the diagnostic methods of choice in evaluating scoliosis in a tertiary care centre.

References

1. Amendt LE, Ause-Ellias KL, Eybers JL, Wadsworth CT, Nielsen DH, Weinstein SL. Validity and reliability testing of the Scoliometer. Phys Ther 1990;70:108–16.
2. Armstrong GD. The intraclass correlation as a measure of intrerrater reliability of subjective judgments. Nurs Res 1981; 30:314–6.
3. Bartko JJ. The intraclass correlation coefficient as a measure of reliability. Psychol Bull 1966;19:3–11.
4. Bland M. An Introduction to Medical Statistics. Oxford: Oxford University Press, 1987:276–93.
5. Bland JM, Altman DG. Measurements in medicine: The analysis of method comparison studies. The Statistician 1983; 32:307–17.
6. Bunnell WP. An objective criterion for scoliosis screening. J Bone Joint Surg [Am] 1984;66:1381–7.
7. Barwell RG, Webb JK, More EJ. School screening for scoliosis (letter). Lancet 1981;2:863.
8. Coughlin SS, Pickle LW, Goodman MT, Wilkens LR. The logistic modeling of interobserver agreement. J Clin Epidemiol 1992;45:1237–41.
9. Diamond GA, Rozanski A, Forrester JA, et al. A model assessing sensitivity and specificity of tests subject to selection bias. J Chron Dis 1986;39:343–55.
10. Diamond GA. Clinical epistemology of sensitivity and specificity. J Clin Epidemiol 1992;45:9–13.
11. Dutkowsky JP, Shearer D, Schepps B, Orton C, Scola F. Radiation exposure to patients receiving routine scoliosis radiography measured at depth in an anthropometric phantom. J Pediatr Orthop 1990;10:532–4.
12. Fleiss JL. Statistical Methods for Rates and Proportions. 2nd ed. New York: John Wiley & Sons, 1981:212–36.
13. Fletcher RH, Fletcher SW, Wagner EH. Prevention. In: Clinical Epidemiology: The Essentials. 2nd ed. Baltimore: Williams & Wilkins, 1988:157–71.
14. Goldberg MS, Poitras B, Mayo N, Labelle H, Bourassa R, Cloutier R. Observer variation in assessing spinal curvature and skeletal development in adolescent idiopathic scoliosis. Spine 1988;13:1371–7.
15. Haas M. Statistical methodology for reliability studies. J Manipulative Physiol Ther 1991;14:119–32.
16. Haas M. The reliability of reliability. J Manipulative Physiol Ther 1991;14:199–208.
17. Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
18. Hanley J, McNeil B. A method of comparing the areas under the receiving operating characteristic curve derived from the same cases. Radiology 1983;148:839–43.
19. Hildreath NG, Shore RE, Woodward E, et al. Risk of extrathyroid tumors following radiation treatment in infancy for thymic enlargement. Radiat Res 1985;102:378–91.
20. Hlatky MA, Pryor DB, Harrell FE, Califf RM, Mark DB, Robery RA. Factors affecting sensitivity and specificity of exercise electrocardiography: Multivariate analysis. Am J Med 1984;77:64–71.
21. Hoffman DA, Lonstein LE, Morin MM, Visscher W, Harris BSH III, Boice JD. Breast cancer in women with scoliosis exposed to multiple diagnostic x rays. J Natl Cancer Inst 1989 81:1307–12.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–77.
23. Levy AR, Goldberg MS, Hanley JA, Mayo NE, Poitras B. Projecting the lifetime risk of cancer from exposure to diagnostic ionizing radiation for adolescent idiopathic scoliosis. Health Phys 1994;66:621–33.
24. Lonstein JE, Carlson MC. Prediction of curve progression in untreated idiopathic scoliosis during growth. J Bone Joint Surg [Am] 1984;66:1061–71.
25. Mausner JS, Bahn AK. Epidemiology: An Introductory Text. 2nd ed. Philadelphia: WB Saunders, 1985:214–38.
26. Mubarak SJ, Wyatt MP, Leach J. Evaluation of the intra-examiner reliability of the scoliometer in measuring trunk rotation. Orthop Trans 1984;9:113–4.
27. Müller R, Büttner P. A critical discussion of intraclass correlation coefficient. Stat Med 1994;13:2465–76.
28. Murrell GAC, Coonrad RW, Moorman CT, Fitch RD. An assessment of the reliability of the scoliometer. Spine 1993;18: 709–12.
29. Nachemson AL, Peterson LE. Effectiveness of treatment with a brace in girls who have adolescent idiopathic scoliosis. J Bone Joint Surg [Am] 1995;77:815–22.
30. Nash CL, Gregg EC, Brown RH, Pillai K. Risks of exposure to x rays in patients undergoing long-term treatment for scoliosis. J Bone Joint Surg [Am] 1979;61:371–4.
31. Pearsall DJ, Reid GJ, Hedden DM. Comparison of three noninvasive methods for measuring scoliosis. Phys Ther 1992; 72:648–57.
32. Streiner DL, Norman GR. Health Measurement Scales. New York: Oxford University Press, 1989:75–95.
33. Thomas AMC. The spine. In: Pynsent P, Fairbank J, Carr A, eds. Outcomes Measures in Orthopedics. Oxford: Butterworth-Heinemann Ltd., 1993:95–103.
34. Tokunaga M, Land CE, Yamamoto T. Incidence of female breast cancer among atomic bomb survivors: Hiroshima and Nagasaki, 1950–1980. Radiat Res 1987;11:243–72.

Address reprint requests to

J. David Cassidy, DC, PhD, FCCSC
Director of Research
Institute for Health and Outcomes Research
PO Box 108, Royal University Hospital
103 Hospital Drive
Saskatoon, Saskatchewan
Canada S7N 0W8