Validity and Reliability of the iPhone to Measure Rib Hump in Scoliosis

Frederic Balg, MD, Mathieu Juteau, Chantal Theoret, MD, Amy Svotelis, PhD, and Guillaume Grenier, PhD

Study Design: This was a prospective blinded validity and reliability analysis.

Objective: The aim of this study was validation and reliability evaluation of the Scoligauge iPhone app.

Background: The scoliometer is used to clinically measure the rib hump in scoliosis as a means to evaluate the axial trunk rotation. The increasing availability of smartphone with built-in accelerometer led to the development of a vast number of applications to measure angles. Of these, the Scoligauge mimics a scoliometer. The aim of this study was to compare the validity of the Scoligauge iPhone application without an associated adapter with the traditional scoliometer and to test the reliability of the application in a clinical setting.

Methods: Two observers measured the rib hump deformity on 34 consecutive patients with idiopathic scoliosis with an average Cobb angle of 24.2 ± 13.5 degrees (range, 4 to 65 degrees). Measurements were made with an iPhone without the adapter and with a scoliometer. The validity as well as the interobserver and intraobserver reliability were calculated using the intraclass coefficient (ICC) and the Bland-Altman test.

Results: The mean difference between the scoliometer and the Scoligauge application was 0.4 degrees [95% confidence interval (CI) of ± 3.1 degrees] with an ICC of 0.947 ($P < 0.001$). The intraobserver and interobserver ICC were 0.961 ($P < 0.001$) and 0.901 ($P < 0.001$), respectively. The mean intraobserver difference was 0.0 degrees (95% CI of ± 2.7 degrees) and the mean interobserver difference was 0.1 degrees (95% CI of ± 4.4 degrees).

Conclusions: The intraobserver and interobserver reliability of the Scoligauge iPhone app, as well as its validity compared with the scoliometer, are excellent. The mean differences between measurements are small and clinically not significant. Thus, the Scoligauge application is valid for clinical evaluation even without special adapter.

Level of Evidence: Level I (Diagnostic Study).

From the Department of Orthopaedic Surgery, Faculty of Medicine and Health Science, University of Sherbrooke, and Centre de recherche du CHUS, Sherbrooke, QC, Canada.

Supported by grant from the Fondation pour la Recherche et l'Enseignement en Orthopédie de Sherbrooke (FREOS) scholarship. M.J. received a scholarship and G.G. received an investigator support from Fonds de la Recherche en Santé du Québec (FRSQ). G.G. and F.B. hold grants from the Canadian Institutes of Health Research (CIHR).

The authors declare no conflicts of interest.

Reprints: Frederic Balg, MD, 3001, 12th Ave. North, Sherbrooke, QC, Canada J1H 5N4. E-mail: frederic.balg@usherbrooke.ca.

Key Words: scoliometer, Scoligauge, rib hump, axial trunk rotation, scoliosis, iPhone, app, smartphone

(*J Pediatr Orthop* 2014;34:774–779)

Evaluation of scoliosis in the frontal plane is measured by x-ray analysis using the Cobb angle. The rotational deformity is best evaluated on computed tomography (CT) scan, but this implies a great amount of radiation for adolescents and is not routinely performed. Rib hump, which can be clinically measured, is the result of a vertebrae rotation. A simple method used to measure the rib hump consists of the forward bending test (the Adam test) coupled with a scoliometer to measure the axial trunk rotation (ATR).[1] This device, which was described by Bunnell, is widely used and has been validated by multiple studies[2–5] and a validated correlation exists between the ATR angle and the Cobb angle.[6,7]

With the arrival of smartphones in the market, many medical applications have been created, such as a goniometer,[8] a heart rate extractor,[9] and a hallux valgus angle calculator.[10] Because smartphones, like the iPhone (Apple Inc., Cupertino, CA), incorporate a micro-electro-mechanical-system (MEMS) accelerometer that can accurately sense acceleration and inclination, several applications were developed to facilitate the measurement of inclination for various purposes. Hence, an application for smartphones (Cobbmeter) was developed for clinicians and validated to measure the Cobb angle in scoliosis.[11,12]

The application called Scoligauge (Ockendon Partners Ltd, UK, http://www.ockendon.net) was created to mimic a scoliometer and measure the ATR with an iPhone. Franko et al[13] performed a biomechanical validation of the application by using an iPhone directly lying on a scoliometer at various random angulations. Izatt et al[14] validated the application on plaster molds from real patients using an adapter to mimic the shape of the scoliometer. Both studies proved the iPhone with the Scoligauge application to be valid against the scoliometer in a controlled nonclinical setting. However, the use of the Scoligauge application with the iPhone without special adapter is still not valid for clinical use.

The aim of this study was to compare the validity of the Scoligauge iPhone application without adapter against the scoliometer and to test the reliability of the application in a clinical setting. For this purpose, we tested the intraobserver and interobserver reliability of the Scoligauge app and tested its validity compared with the scoliometer. We hypothesized that the Scoligauge iPhone application will be a valid instrument that could be used in a clinical setting even without the specific adapter.

MATERIALS AND METHODS

Patients

Thirty-four consecutive patients with adolescent idiopathic scoliosis (AIS) were recruited from a scoliosis clinic. Of this number, 25 were female and 9 were male patients, with ages ranging from 8 to 20 years old, with a mean age of 14.7 ± 2.9 years. All patients displayed measurable scoliosis on x-ray using the Cobb angle ranging from 4 to 65 degrees, with an average of 24.2 ± 13.5 degrees. Nine patients had curvature <15 degrees, 15 were between 15 to 30 degrees, and 10 were >30 degrees. Scoliosis was present in either the thoracic (mean = 27.5 ± 14.5 degrees), thoracolumbar or lumbar (mean = 21.6 ± 12.3 degrees) regions. Patients with neuromuscular or congenital scoliosis were excluded from the study. Patients with severe mental retardation were also exempt. Scoliometer and iPhone measurements were collected after verbal and written consent from patients. This study was approved by the local ethics committee (Protocol #12-132).

Instrumentation

The gold standard measures were made using the original scoliometer (Mizuho OSI, Union City, CA). The electronic measures were made using the Scoligauge application, version 1.3 (Ockendon Partners Ltd, UK, http://www.ockendon.net), installed on iPhone 4S (Apple Inc.) running on iOS 5.1.

Rib Hump Measurement

Two observers measured the ATR among all the patients of this study. The observers consisted of an experienced spinal orthopaedic surgeon (observer #1) and a physical therapy student (observer #2) who acted as the inexperienced observer. Measurement of the rib hump was performed with the Scoligauge iPhone application and with the scoliometer, acting as the gold standard for this study. All measurements were made during the same outpatient visit to avoid asking young patients to miss another school day. To minimize bias, observers were blinded to one another. To assess the interobserver reliability the ATR was measured with the iPhone by the 2 observers at the beginning of the visit. At the end of the visit, the experienced spinal orthopaedic surgeon measured the ATR with the iPhone to assess the intraobserver reliability and with the scoliometer to assess the validity of the Scoligauge app. No measurements were repeated at the end of the visit by observer #2 to decrease patient fatigue. Both observers were asked to place the scoliometer or iPhone at the level they evaluated as the point of maximal rib hump deformation. For double curved scoliosis, measurements of the 2 possible rib humps were performed by the same method. For example, observer #2 measured the 2 ATR specifying the region (either thoracic, thoracolumbar, or lumbar) and then observer #1 made his own measurements. The unit of measure for the scoliometer and Scoligauge app is degrees without decimals. No adapter was used with the iPhone.

Statistical Analysis

The intraclass coefficient (ICC) was used to assess the intraobserver and interobserver reliability and validity. For the intraobserver reliability, the 2 measurements made by the experienced spinal orthopaedic surgeon (observer #1) with the iPhone were compared. The ICC for the interobserver reliability was obtained with the iPhone measurements made by observer #2 and the mean of the 2 smartphone measurements taken by observer #1. The ICC of the Scoligauge application validity was obtained using the mean of the iPhone measurements of observer #1 and the values obtained with the scoliometer. An ICC >0.75 is considered good, over 0.85 very good, and over 0.9 excellent.[15]

The 3 parameters (intraobserver, interobserver reliability, and validity) were also compared using the Bland-Altman plot[16] to compare the difference between each coupled values. Graphically, the mean difference and the 95% confidence interval (CI) were represented. We set the "clinically acceptable" CI at 7 degrees, giving a range of 3.5 degrees above and below the mean of difference.[17] Another criteria for a "clinically acceptable" range was when the mean difference between the values was close to 0.

RESULTS

The 2 observers performed the assessment of the ATR using the rib hump deformation on 34 patients, of which 21 patients had a double deformation measured. This gave a total of 55 deformations and 165 measurements. The mean value of the scoliometer and Scoligauge app were 5.8 degrees (range, 0 to 18 degrees) and 6.2 (range, 0 to 20 degrees), respectively.

Validity

For the totality of the results, the ICC was calculated at 0.947 ($P < 0.001$). For thoracic and lumbar (including thoracolumbar) results, the ICC was 0.965 ($P < 0.001$) and 0.931 ($P < 0.001$), respectively. The Bland-Altman plot (Fig. 1) shows a mean absolute difference of 0.4 degrees with a 95% CI of 3.1 degrees. When divided into specific regions, the thoracic Bland-Altman results illustrate an average difference of 0.3 degrees with a 95% CI of 2.8 degrees. For lumbar results, the mean was 0.4 degrees with an associated 95% CI of 3.3 degrees. For curves <15 degrees, the mean difference was 0.5 degrees (95% CI of 1.5 degrees), for curves between 15 to 30 degrees, the mean difference was 0.0 degrees (95% CI of 3.5 degrees), and for curves >30 degrees, the mean difference was 0.7 degrees (95% CI of 4.0 degrees).

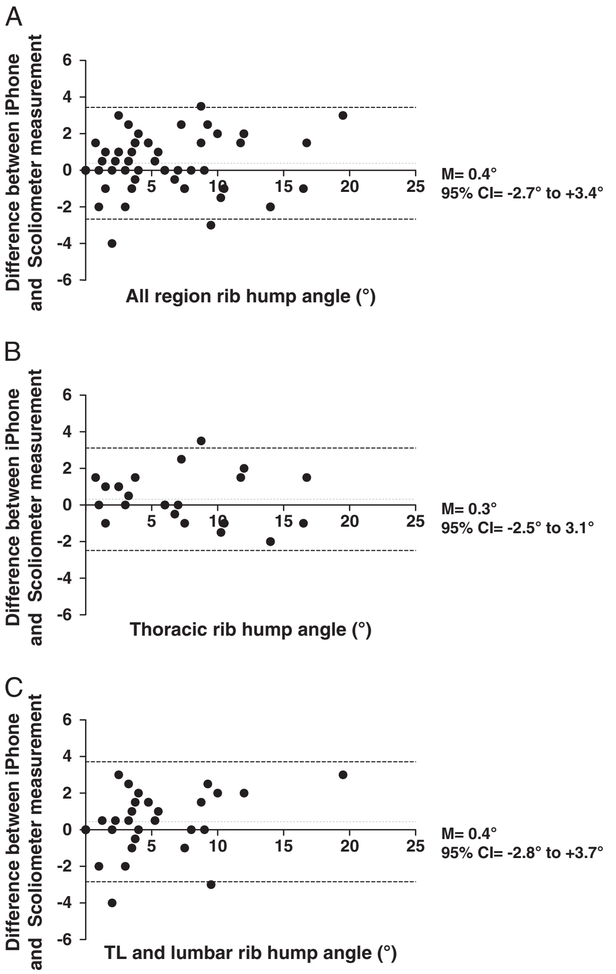

FIGURE 1. The Bland-Altman plot showing the coupled-difference between the Scoligauge iPhone application and scoliometer measurement for (A) all region, (B) thoracic region, and (C) thoracolumbar and lumbar region. Centered gray dotted lines show the mean difference and the 2 outer dotted lines represent the 95% confidence interval (CI).

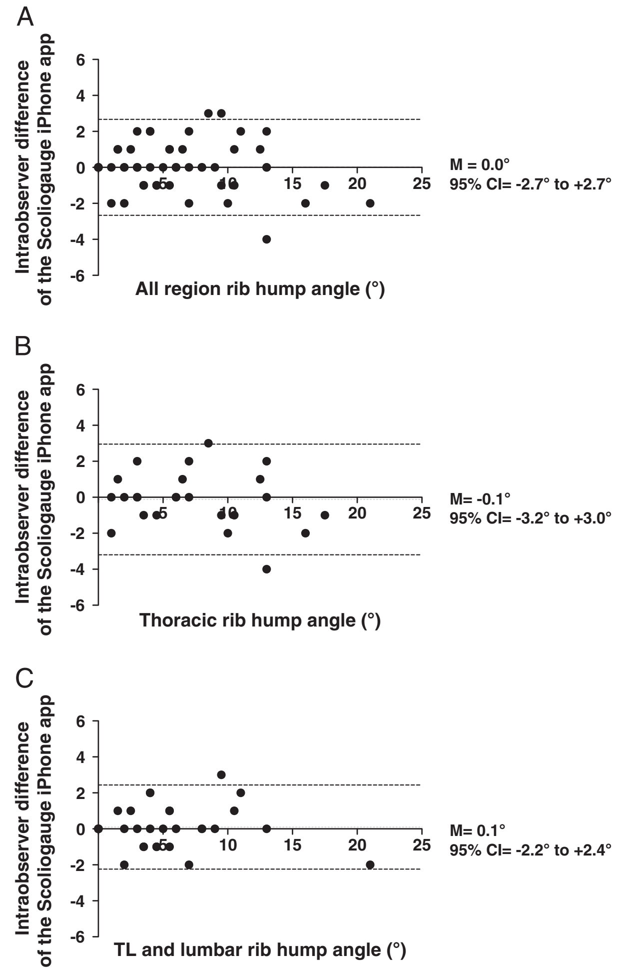

FIGURE 2. The Bland-Altman plot showing the intraobserver coupled-difference of the Scoligauge iPhone application for (A) all regions, (B) thoracic region, and (C) thoracolumbar and lumbar region. Centered gray dotted lines showing the mean difference and the 2 outer dotted lines represent the 95% confidence interval (CI).

Intraobserver Reliability

The intraobserver reliability ICC was 0.961 ($P < 0.001$). Similarly, thoracic measures showed an ICC of 0.952 ($P < 0.001$) and lumbar measurements were calculated at 0.966 ($P < 0.001$). Further, the Bland-Altman plot (Fig. 2) of these measures shows a mean of 0.0 degrees with a 95% CI of 2.7 degrees. Plots for thoracic and lumbar region were similar, with $M = -0.1 \pm 3.1$ degrees and $M = 0.1 \pm 2.3$ degrees, respectively.

Interobserver Reliability

The comparison between these 2 observers gave an ICC of 0.901 ($P < 0.001$). The results for thoracic (0.979; $P < 0.001$) were better than those taken of lumbar regions (0.816; $P < 0.001$). A Bland-Altman plot (Fig. 3) showed a mean absolute difference of 0.1 degrees with a 95% CI of 4.4 degrees. When divided into thoracic and lumbar regions, the results differ in that the mean of thoracic regions was 0.1 degrees with 95% CI of 2.5 degrees and the lumbar mean was 0.2 degrees with a larger 95% CI of 5.5 degrees.

DISCUSSION

This study demonstrated the validity of the Scoligauge application on the iPhone 4S in a clinical setting on scoliosis patients. On the basis of intraobserver and interobserver reliability, we show here that the scoliometer and the Scoligauge are clinically equivalent for measuring rib humps. We also argue that the Scoligauge application on this system can serve as a reliable substitute for the scoliometer without the use of the iPhone adapter.

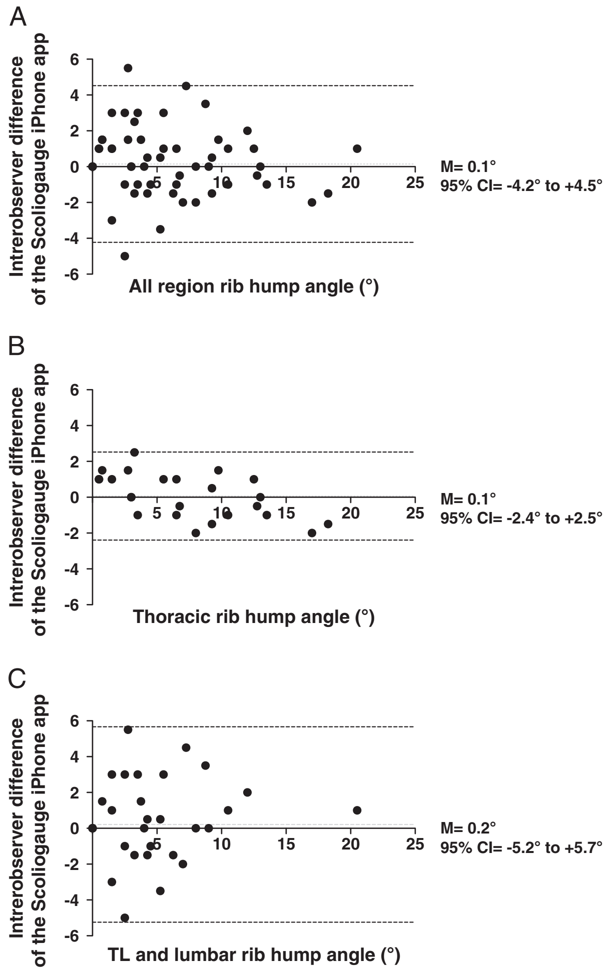

FIGURE 3. The Bland-Altman plot showing the interobserver coupled-difference of the Scoligauge iPhone application for (A) all regions, (B) thoracic region, and (C) thoracolumbar and lumbar region. Centered gray dotted lines showing the mean difference and the 2 outer dotted lines represent the 95% confidence interval (CI).

Our study validated the use of the Scoligauge application without the iPhone adapter, an instrument that mimics the notch provided in the scoliometer to allow the device to fit over the spinous processes. When the 2 devices are compared (scoliometer vs. iPhone), there are some morphologic differences that may affect positioning and proper evaluation of patients. The scoliometer has a length of 16.5 cm, whereas the iPhone 4S that we used has a length of 11.5 cm. This may influence the positioning of the device on patients backs. The scoliometer also has a groove at its center to help clear prominent spinal processes. Another way of positioning the iPhone is to place both thumb between the back of the patient and each side of the under surface of the iPhone to help clear the spinal process as shown on the Scoligauge homepage (http://www.ockendon.net/Scoligauge_Home_Page.htm). Although interesting, in our opinion this may result in a false impression of reproducibility because the thumb is not round but oval. Thumbs placed at different rotation would mean a different space between the iPhone and the patient back, but we did not test this option. Other authors have used an adapter sleeve for the iPhone to better replicate the original shape of the scoliometer.[13,14] Users are required to transport the adapter to the clinic or have one readily available, which may lead to the evaluator being tempted to use it without the adapter. Moreover, the positioning of the iPhone in the adapter is crucial, and an imperfect fit may result in biased measures or additional variability. The use of an adapter adds an interface between the patient and the instrument that could bias the angulation if not properly seated in the adapter. We believe that the availability of these adapters will preclude its use for many nonspecialized clinics and nonspecialists. Of note, the life cycle of electronic devices is short and the shape of new smartphone may render an adapter obsolete. For example, the length of the new iPhone 5 is longer than its predecessor by 8 mm. The iPad could also be used but would require another adapter. For these reasons, we decided to test the validity of the Scoligauge application using an iPhone without adapter. To help position the iPhone on the back of patients with different morphology, both observers were asked to visually align the iPhone to best reflect the true rib hump if the spinal process was too prominent (Fig. 4). Our results include this additional variability. This explains our greater 95% CI.

The study clearly shows that the iPhone is a valid instrument to measure rib hump on patients with AIS. Considering the calculated ICC of 0.947, there is an excellent correlation between the iPhone measurement and the scoliometer measurement. To ensure that there was no bias created by a specific region of the spine, the results were separated in 2 categories: thoracic and lumbar. The specific ICC of these data show that irrespective of the region that is evaluated, the correlation is excellent and we can confirm that measures taken by the iPhone can be used. As shown on the Bland-Altman plot (Fig. 1), we clearly observe the distribution of variability between the 2 methods of measurement. The mean of 0.4 ± 3.1 degrees confirms that the scoliometer and the Scoligauge are clinically equivalent for measuring rib humps. The 95% CI, ranging from -2.7 to $+3.4$ degrees, is an acceptable margin of error equivalent to reported variability for the scoliometer.[2,4,5] The variability between the 2 methods is not clinically significant and therefore allows the use of the Scoligauge iPhone application. In the same vein, thoracic and lumbar regions ICC were examined and there was no bias created by these components.

Franko et al[13] report an excellent correlation of 0.9995 ($P < 0.001$) between the iPhone and scoliometer. As all measurements were taken with the iPhone lying directly on the scoliometer held at different arbitrary angles in the study, this led to potentially more variation in our methodology as we tested the instruments separately on real patients. In our study, observers had to individually place the scoliometer or iPhone on the patient's maximal rib hump deformity and were blinded to the other observer. In the study by Izatt et al,[14] observers

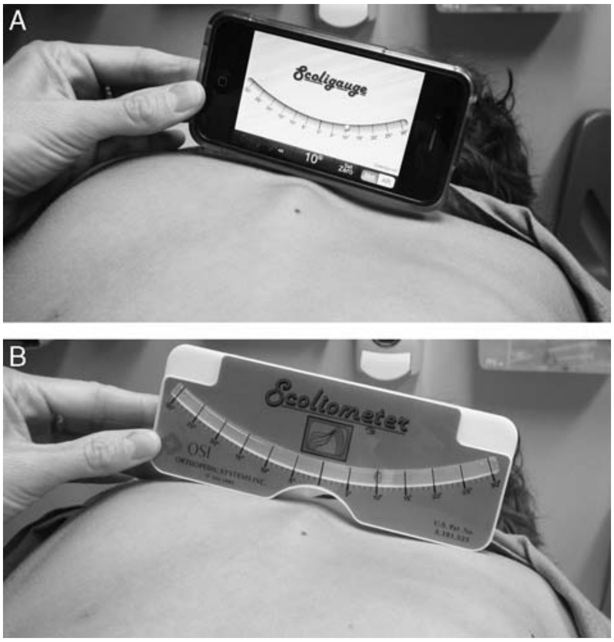

FIGURE 4. Visual alignment of the iPhone to overcome the prominent spinal process and best estimate the true rib hump (A). With the center of the iPhone positioned on the spinous process, the distance between both extremities of the iPhone and the patient are made equal by visual approximation. Same measurements are made with the scoliometer as a gold standard (B).

also had to place the scoliometer or the iPhone themselves on the rib hump and they report an ICC of 0.924, which is consistent with our results. They report a mean difference of 2.1 ± 1.6 degrees between the scoliometer and the iPhone, which is clinically acceptable. Our study shows a smaller mean difference of 0.4 degrees but a larger 95% CI at ± 3.1 degrees. This is still clinically acceptable and can be explained by the difference in methodology. They performed all measurements on plaster molds, which do not have the same deformability like the skin, even if foam is added to the plaster to mimic skin. Moreover, a plaster molding is static, as opposed to real patient examination where they can bend slightly differently between all assessments. We deliberately choose to perform our study in real clinical situation to determine if the Scoligauge iPhone application can be used in the day-to-day clinical settings.

The present study also showed that the smartphone application presents an excellent intraobserver and interobserver reliability (ICC = 0.961 and 0.901, respectively). This compares favorably with the validation of the scoliometer by Bonagamba et al.[3] They obtained ICC of 0.74 to 0.92 for intraobserver reliability and 0.57 to 0.95 for interobserver reliability. When compared with these results, our ICC results were better than the scoliometer, which is commonly used to evaluate rib hump in clinic. This can be explained by the numerical data displayed by the iPhone that does not require interpretation compared with the scoliometer analogic reading. However, the Bland-Altman difference plot shows a large 95% CI, meaning that for the same rib hump, 2 different observers could have a difference of 4.4 degrees. When subdivided into thoracic and lumbar region, we can see that the results for thoracic measurements are more reliable (difference of 2.5 degrees) and that lumbar measurements can create a bias (difference of 5.5 degrees). This result is clinically important as patients consult > 1 health professional, and this can lead to difficulty in following the progression of AIS in these situations. This observation may be because of the fact that thoracic deformation can be easier to assess as it stems from osseous rib deformity, whereas lumbar deformation is mostly due to paravertebral muscle deformation. A large 95% CI is also noted with the use of the scoliometer as reported by Coté et al[4] with a difference of 0.2 ± 4.9 degrees for thoracic measurements and 1.2 ± 6.6 degrees for lumbar measurements. This variability in rib hump measures may be influenced more by the technique of ATR measurement during the forward bending test than the instrument of measure itself.

An interesting use of the Scoligauge application would be in screening of scoliosis before orthopaedic consultation. Although performed in some countries with interesting results,[18–20] the screening of AIS in schools is still controversial, and this practice has been abandoned elsewhere because of a high rate of unnecessary referral.[21,22] If a screening program in the education system is encouraged, it should rely on objective criteria.[6] The scoliometer would have a role to play in such a program, but its cost and availability may impede its use. With the increasing use and availability of smartphones, a low-cost application like Scoligauge could help democratize the evaluation of ATR. However, a major downside of both the scoliometer and the Scoligauge is the CI that maybe too large to be sensitive and specific enough to be used as a screening tool. The use of the Scoligauge as a screening tool is controversial and beyond the scope of this article.

This study proved that even without an adapter, the Scoligauge iPhone application is valid and can be used in the clinical setting for scoliosis evaluation. Although the Scoligauge app with this specific combination is valid, it should be mentioned that the "gold-standard" scoliometer is not constrained by its shape, does not induce additional source of error from the additional interface created by an adapter, is not influenced by software or hardware updates, and will never run out of battery. Although our assessment of this device cannot be generalized to all similar applications, or to different electronic devices, we show that this is a promising avenue for easy and rapid evaluation of rib hum deformation in a clinical setting.

ACKNOWLEDGMENT

The authors thank Nathalie Carrier for her precious help for statistical analysis.

REFERENCES

1. Bunnell WP. An objective criterion for scoliosis screening. *J Bone Joint Surg [Am]*. 1984;66:1381–1387.

2. Amendt LE, Ause-Ellias KL, Eybers JL, et al. Validity and reliability testing of the Scoliometer. *Phys Ther*. 1990;70:108–117.
3. Bonagamba GH, Coelho DM, Oliveira AS. Inter and intra-rater reliability of the scoliometer. *Rev Bras Fisioter*. 2010;14:432–438.
4. Côté P, Kreitz BG, Cassidy JD, et al. A study of the diagnostic accuracy and reliability of the Scoliometer and Adam forward bend test. *Spine*. 1998;23:796–802.
5. Murrell GA, Coonrad RW, Moorman CT, et al. An assessment of the reliability of the Scoliometer. *Spine*. 1993;18:709–712.
6. Bunnell WP. Selective screening for scoliosis. *Clin Orthop Relat Res*. 2005;434:40–45.
7. Sapkas G, Papagelopoulos PJ, Kateros K, et al. Prediction of Cobb angle in idiopathic adolescent scoliosis. *Clin Orthop Relat Res*. 2003;411:32–39.
8. Ferriero G, Sartorio F, Foti C, et al. Reliability of a new application for smartphones (DrGoniometer) for elbow angle measurement. *J Inj Funct Rehabil*. 2011;3:1153–1154.
9. Kwon S, Lee J, Chung GS, et al. Validation of heart rate extraction through an iPhone accelerometer. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society. Boston, Massachusetts USA, August 30 - September 3, 2011. 2011:5260-5263.
10. Ege T, Kose O, Koca K, et al. Use of the iPhone for radiographic evaluation of hallux valgus. *Skeletal Radiol*. 2013;42:269–273.
11. Qiao J, Liu Z, Xu L, et al. Reliability analysis of a smartphone-aided measurement method for the Cobb angle of scoliosis. *J Spinal Disord Tech*. 2012;25:E88–E92.
12. Shaw M, Adam CJ, Izatt MT, et al. Use of the iPhone for Cobb angle measurement in scoliosis. *Eur Spine J*. 2012;21:1062–1068.
13. Franko OI, Bray C, Newton PO. Validation of a scoliometer smartphone app to assess scoliosis. *J Pediatr Orthop*. 2012;32: e72–e75.
14. Izatt MT, Bateman GR, Adam CJ. Evaluation of the iPhone with an acrylic sleeve versus the Scoliometer for rib hump measurement in scoliosis. *Scoliosis*. 2012;7:14.
15. Rosner BA. *Fundamentals of Biostatistics*. Boston, MA: Duxbury Press; 2011.
16. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. *Ultrasound Obstet Gynecol*. 2003;22:85–93.
17. Samuelsson L, Norén L. Trunk rotation in scoliosis. The influence of curve type and direction in 150 children. *Acta Orthop Scand*. 1997;68:273–276.
18. Adobor RD, Riise RB, Sorensen R, et al. Scoliosis detection, patient characteristics, referral patterns and treatment in the absence of a screening program in Norway. *Scoliosis*. 2012;7:18.
19. Adobor RD, Rimeslatten S, Steen H, et al. School screening and point prevalence of adolescent idiopathic scoliosis in 4000 Norwegian children aged 12 years. *Scoliosis*. 2011;6:23.
20. Luk KDK, Lee CF, Cheung KMC, et al. Clinical effectiveness of school screening for adolescent idiopathic scoliosis: a large population-based retrospective cohort study. *Spine*. 2010;35: 1607–1614.
21. Karachalios T, Sofianos J, Roidis N, et al. Ten-year follow-up evaluation of a school screening program for scoliosis. Is the forward-bending test an accurate diagnostic criterion for the screening of scoliosis? *Spine*. 1999;24:2318–2324.
22. Plaszewski M, Nowobilski R, Kowalski P, et al. Screening for scoliosis: different countries; perspectives and evidence-based health care. *Int J Rehab Res*. 2012;35:13–19.