Inter- and Intra-rater Reliability of Postural Assessment for Scoliosis

Chompunoot Suwanasri MSc*,
Wimonrat Sakullertphasuk MSc*, Meena Tosiriphattana BSc*,
Thachakorn Sa-ngounsak BSc*, Worawan Ekabutr BSc*

*Faculty of Physical Therapy, Mahidol University, Nakhon Pathom, Thailand

Objective: *The purpose of the present study is to assess the inter- and intra-rater reliability of a postural assessment for detecting scoliosis between 4th year physical therapy (PT) students and PT specialists by using 30 postural indices.*
Material and Method: *Six examiners, three 4th year PT students and 3 certified PT specialists, performed the postural indices on 10 asymptomatic subjects.*
Results: *Inter-rater reliability between 4th year students and PT specialists and intra-rater reliability ranged from poor to almost perfect. Two items that needed rectification were PSIS level and iliac crest.*
Conclusion: *There was a variable range of values in agreement either within or between examiners for the assessment of scoliosis screening. This shows that the assessment in scoliosis screening should be used with caution by 4th year students.*

Keywords: *Scoliosis, posture assessment, postural index*

J Med Assoc Thai 2014; 97 (Suppl. 7): S1-S5
Full text. e-Journal: *http://www.jmatonline.com*

Scoliosis is a complex three-dimensional deformity of the spine. It can be categorized by a lateral deviation and axial rotation[1,2]. The diagnosis often comes as a complete surprise because parents tend not to see their children's bare back[3]. It is important that scoliosis be accurately diagnosed to determine an appropriate treatment, and to prevent progressive scoliosis. Physical therapists (PTs) apply a variety of procedures to detect scoliosis, plan treatment and prevent the progression of scoliosis. The postural assessment for scoliosis consists of visual observation, palpation, and motion testing[2]. The previous studies[4,5] of postural assessment were carried out by 4th year physical therapy (PT) students using several indices. However, the indices of assessments were not the same and the investigators did not examine the inter- and intra-reliability of measurements. Therefore, the present study was to develop the indices and to assess the inter- and intra-rater reliability of postural assessment for scoliosis between 4th year PT students and PT specialists.

Correspondence to:
Suwanasri C, Faculty of Physical Therapy, Mahidol University 999 Phuttamonthon 4 Road, Salaya, Nakhon Pathom 73170, Thailand.
Phone: 0-2441-5450 ext. 21003, Fax: 0-2441-5454
E-mail: chompunoot.suw@mahidol.ac.th

Material and Method

The present study was approved by Mahidol University Institutional Review Board (MU-IRB 2013/035.2504). The examiners consisted of three 4th year PT students and three PT specialists training for scoliosis management. The subjects were recruited from the undergraduate students at Mahidol University. Inclusion criteria were as follows: the body mass index (BMI) less than 23 kg/m^2, no pain, no limitation of any joints and able to stand still longer than 15 minutes. The subjects unable to perform forward bending test were excluded.

Procedure

The researchers developed 30 indices of postural assessment for scoliosis and guidelines on definition of each index. Such indices were examined in anterior-posterior and lateral views, and the forward bending test; index 1-19, all of the examiners assessed from a posterior view, index 20-24 from lateral view, index 25-28 from anterior view, and index 29 to 30 from forward bending test. In some indices, the examiners assessed by using a palpation for comparing the level of bony landmark, for instance, index 8, 11 to 13, and 23. Details of indices were illustrated under table 1 to 4. The subjects stood on the step that had been set up as shown in Fig. 1. The subjects wore shorts with the

length above their ankles. Male subjects were asked to take off their shirt while females were asked to wear the backless shirt, which was prepared for observing the line of spinal curvature clearly. All examiners assessed subjects by observation, palpation and judgment for 10 minutes and recorded the result in the scoliosis assessment form. PT specialists assessed each participant once while the 4th year students assessed each participant once a day and assessed three times per each participant (with more than one-day interval between each assessment for preventing the students' recognition).

Data analysis

The statistical analysis was performed by Cohen's kappa coefficient (K), which was calculated by using Statistical Package for Social Science (SPSS) version 21.0. Inter-rater reliability was computed by using an agreement from the certified PT specialists, which calculated from 2 out of 3 agreements, and compared with the 4th year PT students. This study used three specialists for increasing the accuracy of results as well as creating standard agreement in the assessment process. Intra-rater reliability of each student was calculated by comparing the agreements at 1st and 2nd trials from three assessment trials of each participant. Kappa values above zero were interpreted using the scale purposed by Landis and Koch, which are 0.0-0.2 = slight, 0.21-0.40 = fair, 0.41-0.60 = moderate, 0.61-0.80 = substantial, and 0.81-1.00 = almost perfect[6].

Results

Ten subjects (BMI = 19.05 kg/m^2, age = 20$\pm$1 years on average, three males and seven females, were recruited from students at Mahidol University. To distinguish clearly between different result levels, Cohen's Kappa Coefficient were divided into 3 ranges for inter-rater reliability, which are <0.4, 0.4-0.6, and >0.6. Nine of thirty indices in inter-rater reliability ranged from >0.4-1.0. While, five posture indices were shown in Kappa <0.0-0.40, as presented in Table 1. For intra-rater reliability ranged from poor to almost perfect, three indices ranged from moderate to almost perfect, and seven posture indices ranged from poor to fair, as shown in Table 2.

Discussion

Nine of thirty indices in inter-rater reliability ranged from >0.4-1.0, which are shoulder level, scoliosis cervical, scoliosis thoracic, scoliosis lumbar, rib cage, waist form, standing posture (Lt), thoracic prominence, and lumbar prominence. Those indices were not complicated being observed anatomically since few rotations or small changes present. While, five posture indices (popliteal line level, PSIS level, Pelvic level (iliac crest), ASIS level, and pelvis rotation) were shown in <0.4 because of the region of pelvis could not only be observed visually but touching was also needed for the better precision, and these indices were difficult to be assessed and required more skills and experiences in palpation. In terms of popliteal levels, this is because it is not properly viewed from eye level, so this made it harder to assess.

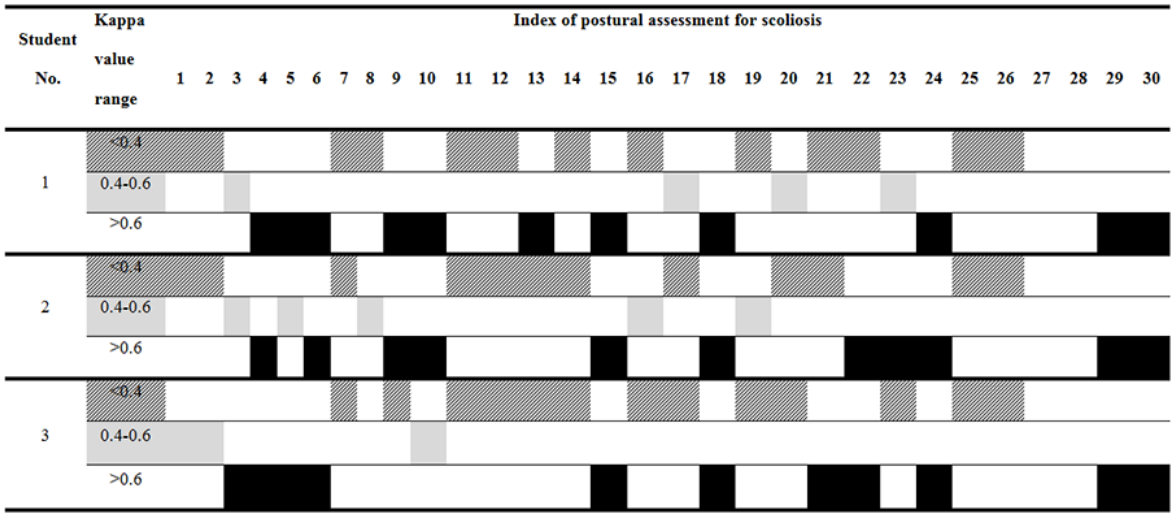
Fig. 1 Dress and position of participants.

Intra-rater reliability ranged from moderate to almost perfect in three of thirty indices (scoliosis cervical, knee alignment, and lumbar prominence), because of those indices being easy to observe anatomically. While, seven postures indices (scapula border, waist form, PSIS level, iliac crest, gluteal fold level, right calcaneal line, and thoracic prominence) ranged from poor to fair because these areas were more difficult to be determined and required more skills and experience.

In inter- and intra-rater reliability, there were some indices that were ventral prominent and showed poor posture, and could not be calculated by using SPSS because those indices lacked data distribution (the examiners recorded the same code for each subject). Therefore, this could be improved by using other statistical methods in the future study.

Error in judgments of posture indices in these experiments may have occurred in the observation, in palpation or in both. The present results may be influenced by the subjects having to change the position, the force used differed between examiners,

Table 1. Inter-rater reliability between 4th year physical therapy students and physical therapy specialists in sub-group



1 = head lateral flexion; 2 = head rotation; 3 = shoulder level; 4, 5, 6 = scoliosis C-T-L; 7 = scapular (border); 8 = inferior angle of scapular; 9 = rib cage (hump); 10 = waist form; 11 = pelvis level (iliac crest); 12 = PSIS level; 13 = gluteal fold level; 14 = popliteal line level; 15 = knee alignment; 16, 17 = Lt & Rt calcaneal line; 18, 19 = standing posture (Lt & Rt); 20 = head; 21, 22 = shoulder in lateral view (Lt & Rt); 23 = pelvic tilt; 24 = anterior chest; 25 = ASIS level; 26 = pelvis rotation; 27 = ventral prominence; 28 = poor posture; 29, 30 = T&L prominence

Table 2. Intra-rater reliability between 4th year physical therapy students

	Intra-rater reliability; Kappa value range: Index of posture					
	<0 Poor	0.0-0.2 Slight	0.21-0.40 Fair	0.41-0.60 Moderate	0.61-0.80 Substantial	0.81-1.0 Almost perfect
ST1	6, 9, 20	7, 11, 13, 16, 19, 25, 26	3, 5, 12, 14, 17, 18	10	1, 8, 29	2, 4, 15, 21-24, 27, 28, 30
ST2	9	11, 12, 18	14, 16, 19, 26	1, 3, 5, 8, 10, 17, 25	2, 6, 13, 20, 29, 30	4, 7, 15, 21-24, 27, 28
ST3	7, 9-12, 21	17, 25	2, 6, 13, 16, 18, 19, 22, 26, 30	1, 5, 8, 14, 20, 29	3	4, 15, 23, 24, 27, 28

1 = head lateral flexion; 2 = head rotation; 3 = shoulder level; 4, 5, 6 = scoliosis C-T-L; 7 = scapular (border); 8 = inferior angle of scapular; 9 = rib cage (hump); 10 = waist form; 11 = pelvis level (iliac crest); 12 = PSIS level; 13 = gluteal fold level; 14 = popliteal line level; 15 = knee alignment; 16, 17 = Lt & Rt calcaneal line; 18, 19 = standing posture (Lt & Rt); 20 = head; 21, 22 = shoulder in lateral view (Lt & Rt); 23 = pelvic tilt; 24 = anterior chest; 25 = ASIS level; 26 = pelvis rotation; 27 = ventral prominence; 28 = poor posture; 29, 30 = T&L prominence; ST = student

variation in examination techniques, and the short amounts of practices of the 4th year PT students, which were just only about 18 hours before the assessment. Two indices, which showed poor to fair agreements in, inter and intra-rater reliability rater reliability, were PSIS level and iliac crest; these indices were difficult to be palpated.

In the present study, it shows less agreement in inter and intra-rater reliability for these two indices, the PSIS level and iliac crest. It was similar to previous studies, which indicated poor reliability in regions of pelvis[7,8]. The result demonstrated that the PSIS level was more difficult to find the exact percent agreement than the iliac crest. While other studies reported the inter-rater reliability of palpation were significantly greater for the PSIS level than the iliac crest, $p<0.05$[9]. This could be an area for a future study. Inter and Intra-rater reliability could possibly be improved by the better

training of examiners[10] and should include a formal or informal comparison of results between examiners followed by several retests until obtaining better agreements.

Conclusion

There was a variable range of value agreements either within or between examiners for the assessment in scoliosis screening. This shows that the general lack of reliability for this scoliosis screening indicating that students should exercise caution when interpreting the assessment.

There were two items for which rectification was suggested, the PSIS levels and the iliac crest. As observation and palpation play an important role in assessment and diagnosis in scoliosis screening program, the variable range of inter and intra-examiner reliability should be a factor to be carefully observed and assessed.

What is already known on this topic?

Scoliosis in Thailand was usually detected through school screening program by fourth year physical therapy students. The basic way to assess scoliosis is the postural assessment using visual observation, static palpation, and motion testing. There is no study dedicated to finding reliability of scoliosis screening program between the physical therapy students and physical therapy specialists.

What this study adds?

The present study shows the general lack of reliability for scoliosis screening indicating that students should practice caution when interpreting the assessment. This suggests that inter- and intra-rater reliability could possibly be improved by the better training of examiners and should have a formal or informal comparison of results between examiners followed by several retests until receiving better agreements.

Potential conflicts of interest

None.

References

1. Charles YP, Dimeglio A, Marcoul M, Bourgin JF, Marcoul A, Bozonnat MC. Influence of idiopathic scoliosis on three-dimensional thoracic growth. Spine 2008; 33: 1209-18.
2. Adobor RD, Riise RB, Sorensen R, Kibsgard TJ, Steen H, Brox JI. Scoliosis detection, patient characteristics, referral patterns and treatment in the absence of a screening program in Norway. Scoliosis 2012; 7: 18.
3. Gardner A. Clinical assessment of scoliosis. Orthop Trauma 2011; 25: 397-402.
4. Meechai P, Intakochasarn S, Boon-im A. Prevalence of various abnormalities which trend to cause scoliosis. Special project of Faculty of Physical Therapy, Mahidol University. Bangkok: Mahidol University; 1997.
5. Nillerd C, Marcharoen V, Kamonwat N. Relationship between weight distribution and spinal alignment in early adolescents. Special project of Faculty of Physical Therapy, Mahidol University. Bangkok: Mahidol University; 2011.
6. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-74.
7. Holmgren U, Waling K. Inter-examiner reliability of four static palpation tests used for assessing pelvic dysfunction. Man Ther 2008; 13: 50-6.
8. Rajendran D, Gallagher D. The assessment of pelvic landmarks using palpation: A reliability study of undergraduate students. Int J Osteopath Med 2011; 14: 57-60.
9. Kim HW, Ko YJ, Rhee WI, Lee JS, Lim JE, Lee SJ, et al. Interexaminer reliability and accuracy of posterior superior iliac spine and iliac crest palpation for spinal level estimations. J Manipulative Physiol Ther 2007; 30: 386-9.
10. Fryer G, McPherson HC, O'Keefe P. The effect of training on the inter-examiner and intra-examiner reliability of the seated flexion test and assessment of pelvic anatomical landmarks with palpation. Int J Osteopath Med 2005; 8: 131-8.

ความน่าเชื่อถือของการตรวจการทรงท่าสำหรับภาวะกระดูกสันหลังคด

ชมพูนุท สุวรรณศรี, วิมลรัตน์ สกุลเลิศผาสุข, มีนา โตศิริพัฒนา, ทัชชกร สงวนศักดิ์, วรวรรณ เอกบุตร

วัตถุประสงค์: *วัตถุประสงค์ของการศึกษาครั้งนี้ เพื่อทดสอบความน่าเชื่อถือ (inter-rater reliability) การตรวจการทรงท่าสำหรับแยกผู้มีกระดูกสันหลังคด โดยนักศึกษากายภาพบำบัดชั้นปีที่ 4 และ ผู้เชี่ยวชาญทางกายภาพบำบัดกระดูกสันหลังคดและภายในผู้วัดเอง (intra- rater reliability) โดยใช้ 30 ดัชนีวัดการทรงท่า*

วัสดุและวิธีการ: *ผู้เข้าร่วมศึกษาทั้งหมด 6 ราย ได้แก่ นักศึกษากายภาพบำบัดชั้นปีที่ 4 จำนวน 3 ราย กับ นักกายภาพบำบัดเชี่ยวชาญทางกระดูกสันหลังคด จำนวน 3 ราย ตรวจประเมิน 30 ดัชนีการทรงท่าในผู้ที่ไม่มีอาการ 10 ราย*

ผลการศึกษา: *ความน่าเชื่อถือระหว่างนักศึกษากายภาพบำบัดชั้นปีที่ 4 กับ นักกายภาพบำบัดเชี่ยวชาญทางกระดูกสันหลังคด และความน่าเชื่อถือการวัดซ้ำของนักศึกษากายภาพบำบัดชั้นปีที่ 4 อยู่ในช่วง ควรปรับปรุงถึงดีมาก มี 2 ดัชนี การทรงท่าที่ควรต้องทำความเข้าใจ ในการตรวจวัด คือ ระดับ PSIS และ iliac crest*

สรุป: *ค่าความน่าเชื่อถือของการตรวจประเมิน ระหว่างนักศึกษากายภาพบำบัดและนักกายภาพบำบัดผู้เชี่ยวชาญ และค่าความน่าเชื่อถือการวัดซ้ำของนักศึกษาปีที่ 4 แสดงค่าในช่วงระดับที่แตกต่างกัน การศึกษาครั้งนี้แสดงให้เห็นว่าในกระบวนการคัดกรองกระดูกสันหลังคดโดยนักศึกษาชั้นปี 4 ควรทำการตรวจประเมินด้วยความระมัดระวัง*