

RESEARCH

Open Access



# Reproducibility of thoracic kyphosis measurements in patients with adolescent idiopathic scoliosis

Søren Ohrtnissen<sup>1\*</sup> , Jason Pui Yin Cheung<sup>2</sup>, Dennis Winge Hallager<sup>1</sup>, Martin Gehrchen<sup>1</sup>, Kenny Kwan<sup>2</sup>, Benny Dahl<sup>1</sup>, Kenneth M. C. Cheung<sup>2</sup> and Dino Samartzis<sup>2\*</sup>

## Abstract

**Background:** Current surgical treatment for adolescent idiopathic scoliosis (AIS) involves correction in both the coronal and sagittal plane, and thorough assessment of these parameters is essential for evaluation of surgical results. However, various definitions of thoracic kyphosis (TK) have been proposed, and the intra- and inter-rater reproducibility of these measures has not been determined. As such, the purpose of the current study was to determine the intra- and inter-rater reproducibility of several TK measurements used in the assessment of AIS.

**Methods:** Twenty patients (90% females) surgically treated for AIS with alternate-level pedicle screw fixation were included in the study. Three raters independently evaluated pre- and postoperative standing lateral plain radiographs. For each radiograph, several definitions of TK were measured as well as L1–S1 and nonfixed lumbar lordosis. All variables were measured twice 14 days apart, and a mixed effects model was used to determine the repeatability coefficient (RC), which is a measure of the agreement between repeated measurements. Also, the intra- and inter-rater intra-class correlation coefficient (ICC) was determined as a measure of reliability.

**Results:** Preoperative median Cobb angle was 58° (range 41°–86°), and median surgical curve correction was 68% (range 49–87%). Overall intra-rater RC was highest for T2–T12 and nonfixed TK (11°) and lowest for T4–T12 and T5–T12 (8°). Inter-rater RC was highest for T1–T12, T1-nonfixed, and nonfixed TK (13°) and lowest for T5–T12 (9°). Agreement varied substantially between pre- and postoperative radiographs. Inter-rater ICC was highest for T4–T12 (0.92; 95% CI 0.88–0.95) and T5–T12 (0.92; 95% CI 0.88–0.95) and lowest for T1-nonfixed (0.80; 95% CI 0.72–0.88).

**Conclusions:** Considerable variation for all TK measurements was noted. Intra- and inter-rater reproducibility was best for T4–T12 and T5–T12. Future studies should consider adopting a relevant minimum difference as a limit for true change in TK.

**Keywords:** Adolescent idiopathic scoliosis, Thoracic, Kyphosis, Radiograph, Sagittal, Flexibility, Reproducibility, Reliability, Agreement, Intra-class correlation, Mixed effects model, Repeatability coefficient, Limits of agreement

## Background

Adolescent idiopathic scoliosis (AIS) is characterized by a lateral deviation of the spine in the coronal plane, vertebral rotation in the transverse plane, and often hypokyphosis in the sagittal plane [1, 2]. Current surgical treatment for AIS involves multisegmental pedicle screw

instrumentation, which results in considerable correction in the coronal plane with limited loss of correction over time [3, 4]. However, several studies have reported failure to restore the thoracic kyphosis (TK) to a normal range seen in non-scoliotic subjects, and in recent years, the importance of surgical correction of sagittal malalignment has gained increased focus [5, 6].

Although measuring TK in AIS patients on plain radiographs has become commonplace throughout the decades, considerable variation across studies in terms of defining TK exists and no consensus has been established on what

\* Correspondence: ohrtnissen@gmail.com; dsamartzis@msn.com

<sup>1</sup>Spine Unit, Department of Orthopedic Surgery, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, Copenhagen East 2100, Denmark

<sup>2</sup>Department of Orthopedics and Traumatology, The University of Hong Kong, Professorial Block, 5th Floor 102 Pokfulam Road, Hong Kong, SAR, China

should be regarded as an actual change in TK as opposed to expected measurement variation. For one, a recent meta-analysis evaluated the surgical correction of TK in AIS patients; however, the analysis included various studies with different definitions of TK, which made direct comparisons challenging [7]. Moreover, several studies have attempted to define the TK range in normal subjects but have used different definitions without addressing differences in measurement variation [8–10]. Furthermore, the Lenke classification [11] is widely used in preoperative planning; however, classification of the sagittal thoracic modifier has shown poor reliability and the measurement agreement for T2–T12 and T5–T12 kyphosis have been found to be reduced compared to the frontal Cobb angle [12, 13]. For T2–T5 regional kyphosis, reliability has been shown to be poor [13, 14] and other studies have shown that the upper part of the thoracic spine is inherently challenging to visualize due to structural overlap of the shoulder girdle [8, 15, 16]. As the clinical importance of the spinal sagittal profile becomes increasingly evident, there is a need to ensure that the measuring methods used to evaluate TK are both accurate and reproducible, especially since the rotational component of the curve may alter reproducibility depending on definitions of TK. Traditionally, TK is determined by a fixed limit Cobb technique (fixed TK, e.g., T4–T12); conversely, the definitions of fixed TK vary among studies [17–20]. A few authors have suggested applying a nonfixed approach where limits of TK are based on the individual sagittal shape of the spine as it has been shown that the cranial and caudal end vertebrae of the non-fixed TK vary among normal adolescents [8, 21–23].

Overall, the intra- and inter-rater reproducibility of these various TK measurements has not been established, and there is no consensus as to which measurements offer the least amount of variability. While a few studies have addressed the intra- and inter-rater correlation for certain TK measurements [24], it is of limited application on individual patients and it will be of great clinical and academic value to know the actual expected variation for repeated TK measurements on the same subject. As such, the objective of the following study was to determine intra- and inter-rater reproducibility of commonly used TK measurements.

## Methods

Plain radiographs of 20 patients who were at one point diagnosed with AIS and surgically treated at our institution with alternate-level pedicle screw fixation [25, 26] were examined. Institutional review board approval was obtained. Gender and patient age was recorded, and curve type was determined based on the Lenke classification [27]. On the coronal radiograph, pre- and postoperative main Cobb angle was measured and correction rate was calculated.

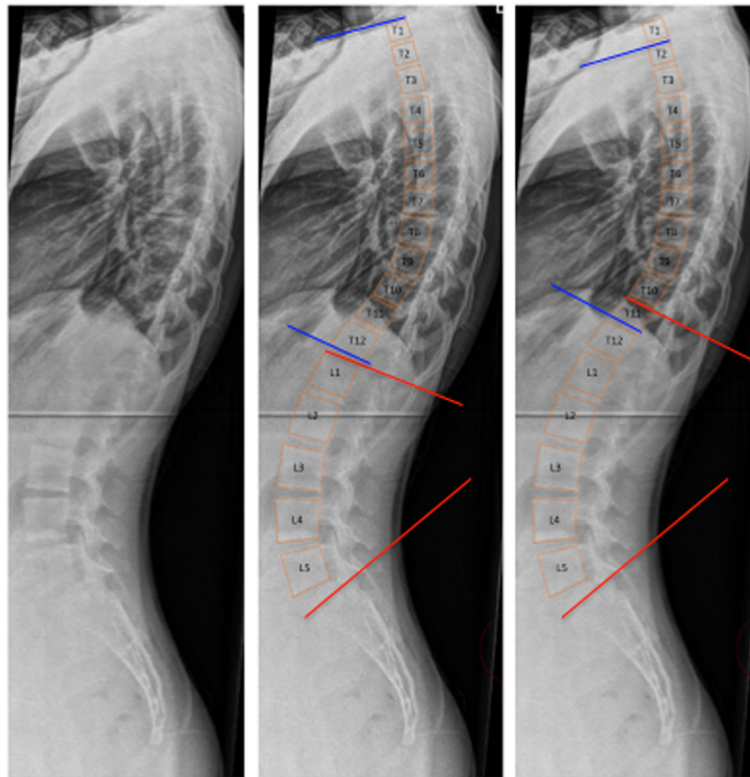
One spine research fellow (rater 1) and two spine surgeons (raters 2 and 3) independently evaluated 20 sets of pre- and postoperative standing lateral radiographs. For each radiograph, the following were determined (Fig. 1):

1. Fixed TK defined as the Cobb angle between the superior end plate and the inferior end plate of T1–T12, T2–T12, T4–T12, and T5–T12 [28]
2. T1-nonfixed TK: From the superior end plate of T1 to the inferior end plate of the most tilted vertebra in the thoracolumbar region [21, 22]
3. Nonfixed TK: From the superior end plate of the most tilted vertebra in the proximal thoracic region to the inferior end plate of the most tilted vertebra in the thoracolumbar region [8]
4. Fixed lumbar lordosis (LL): From the superior end plate of L1 to the superior end plate of S1 [29]
5. Nonfixed LL: From the superior end plate of the most tilted vertebra in the thoracolumbar region to the superior end plate of S1 [23]

Each rater independently performed all measurements twice 14 days apart. Before the second round of measurements, the sequence of the radiographs was randomly reassigned and the raters were blinded from the results of the first round. All raters were blinded to patient details. The total analysis produced 1920 data points for further analysis. All radiographs were measured on a high-resolution monitor using the Picture Archiving Communication system, and identification and labeling of individual vertebrae was based on the Radiographic Measurement Manual by the Spine Deformity Study Group [28]. Application of this manual was discussed among the raters, and consensus was established prior to the study. The protocol for the study was based on the Guidelines for Reporting Reliability and Agreement studies [30].

## Imaging details

For the scoliosis radiographs, all patients were positioned in erect position with the feet together and in the straightest posture possible. For lateral images, patients were in the clavicle position with flexed shoulders and elbows past 90° with hands pointing at the sternal notch to allow better spine visualization while preventing changes to the sagittal balance [31]. A computed detector was utilized to determine the position of the patient's skull and hip joints and also the length of the image required. The detector was 40 cm in length, and thus, image splitting was required. Up to 2–3 exposures were required depending on the patient's height. The postero-anterior radiographs were taken with 78-peak kilovoltage and 20 mAs of X-ray energy. The lateral radiographs were taken with 88-peak kilovoltage and



**Fig. 1** Left: Standing sagittal radiograph of a thoracic single curve with apex at T8. Middle: Fixed measurements of T1–T12 thoracic kyphosis (blue) and L1–S1 (red). Right: Nonfixed measurements of thoracic kyphosis (blue) and lumbar lordosis (red)

32 mAs of X-ray energy. For both images, the focus film distance was 180 cm.

### Statistical analysis

All statistical analyses were performed using R version 3.2.3 (R core team, 2014, Vienna, Austria). Data was reported as proportions (%), mean with standard deviation (SD), or median with range, and data distribution was assessed by histograms.

Reproducibility is a term that entails both measurement agreement and reliability. Intra- and inter-rater agreement is defined as the degree to which repeated measurements are identical whereas reliability is defined as the ability of a measurement to differentiate between subjects [30]. Intra- and inter-rater agreement per subject was estimated for each type of TK measurement using the repeatability coefficient (RC), which is the difference in measurements exceeded by only 5% of pairs of measurements on the same subject. Ninety-five percent limits of agreement were defined as  $\pm RC$ , meaning that a high RC indicated a high variation (poor agreement) in repeated measurements.

Intra-rater agreement for each rater was calculated according to Bland and Altman [32]:

Single rater  $RC = 1.96 * SD$  of the difference between repeated measurements for each rater.

Overall, intra- and inter-rater RC was calculated using a linear mixed effects model with subjects and rater-within-subject variation as random effects and timing of radiograph (e.g., pre- or postoperative) as a fixed effect: [24]

Overall intra-rater  $RC = 2.77 * \sqrt{(\text{residual mean square})}$

Overall inter-rater  $RC = 2.77 * \sqrt{(\text{rater:subject mean square} + \text{residual mean square})}$

Inter-rater RC was further analyzed for pre- and postoperative radiographs separately.

Intra- and inter-rater reliability was estimated with intra-class correlation coefficient (ICC) with 95% confidence interval (CI). We considered an ICC of 0.0–0.24 to represent absent to poor, 0.25–0.49 low, 0.50–0.69 fair/moderate, 0.70–0.89 good, and 0.90–1.0 excellent reliability [33, 34].

### Results

Eighteen patients were female (90%), and the median age was 13.8 years (range 11.5–27.6 years). Eighty-five percent of curves were Lenke type 1 and 15% Lenke type

3, and the preoperative median coronal Cobb angle was 58° (range 41°–86°), which was corrected to a postoperative median Cobb angle of 20° (range 8°–27°) corresponding to a median curve correction of 68% (range 49–87%). Median number of fused levels was 9 (range 6–11 levels). The upper instrumented vertebra was T4, T5, T6, and T7 in one, 13, five, and one patient, respectively. Lowest instrumented vertebra (LIV) was T11, T12, L1, L2, and L3 in two, five, eight, three, and two patients, respectively. Summary of all measurements of both pre- and postoperative radiographs for each round is listed in Table 1 (Additional file 1).

#### Intra- and inter-rater agreement

Single rater RC showed substantial differences among raters ranging from 5° to 13° (Table 2). Overall intra-rater RC was highest for T2–T12, T1-nonfixed, and non-fixed TK (11°) and lowest for T4–T12 and T5–T12 (8°). The overall inter-rater RC was highest for T1–T12, T1-nonfixed, and nonfixed TK (13°) and lowest for T5–T12

(9°) (Table 3). Inter-rater RC ranged between 7° and 14° across pre- and postoperative radiographs. For fixed LL and nonfixed LL, variation was similar to intra- and inter-rater RC ranging from 10° to 11° (Tables 2 and 3).

#### Intra- and inter-rater reliability

Intra-rater ICC was highest for T4–T12 (0.94; 95% CI 0.92–0.96) and T5–T12 (0.94; 95% CI 0.91–0.96) and lowest for T2–T12 (0.84; 95% CI 0.79–0.85) (Fig. 2). Inter-rater ICC was highest for T4–T12 (0.92; 95% CI 0.88–0.95) and T5–T12 (0.92; 95% CI 0.88–0.95) and lowest for T1-nonfixed (0.80, 95% CI 0.72–0.88) (Fig. 3).

#### Discussion

Our study noted a substantial measurement variation for all definitions of TK with the best reproducibility for T4–T12 and T5–T12 both in terms of intra- and inter-rater agreement as well as reliability. Only a few previous studies have addressed the variation of TK measurements in a systematic manner. For example, in a study by Ilharreborde

**Table 1** Summary of measurements for each rater for pre- and postoperative sagittal radiographs of both rounds of measurements

Variable	Rater 1 Cobb angle, mean ± SD		Rater 2 Cobb angle, mean ± SD		Rater 3 Cobb angle, mean ± SD	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
T1–T12						
Preoperative	29 ± 11	29 ± 11	30 ± 11	32 ± 10	34 ± 9	31 ± 10
Postoperative	26 ± 7	26 ± 8	30 ± 9	32 ± 8	31 ± 6	31 ± 7
T2–T12						
Preoperative	28 ± 12	28 ± 12	31 ± 13	30 ± 10	33 ± 14	28 ± 11
Postoperative	26 ± 8	26 ± 9	29 ± 8	31 ± 8	28 ± 6	28 ± 8
T4–T12						
Preoperative	24 ± 13	23 ± 12	26 ± 13	25 ± 12	29 ± 15	26 ± 13
Postoperative	19 ± 7	18 ± 6	21 ± 8	19 ± 7	21 ± 6	20 ± 6
T5–T12						
Preoperative	23 ± 12	21 ± 12	25 ± 12	23 ± 13	25 ± 14	24 ± 14
Postoperative	16 (6)	16 ± 7	17 ± 7	16 ± 6	18 ± 6	18 ± 6
T1-nonfixed						
Preoperative	30 ± 12	30 ± 11	30 ± 12	32 ± 10	35 ± 10	32 ± 11
Postoperative	27 ± 7	27 ± 8	31 ± 9	32 ± 7	32 ± 7	31 ± 7
Nonfixed TK						
Preoperative	32 ± 11	31 ± 11	35 ± 12	35 ± 10	37 ± 11	34 ± 13
Postoperative	28 ± 8	28 ± 9	32 ± 10	34 ± 9	33 ± 7	32 ± 8
L1–S1						
Preoperative	58 ± 11	56 ± 10	55 ± 11	58 ± 12	55 ± 10	56 ± 10
Postoperative	55 ± 8	53 ± 9	51 ± 9	53 ± 9	49 ± 9	52 ± 8
Nonfixed LL						
Preoperative	60 ± 11	60 ± 11	60 ± 12	62 ± 12	58 ± 12	59 ± 11
Postoperative	57 ± 9	55 ± 9	55 ± 9	57 ± 9	52 ± 9	54 ± 9

SD standard deviation, TK thoracic kyphosis, LL lumbar lordosis

**Table 2** Single rater RC for all three raters and overall intra-rater RC with pre- and postoperative subgroups

Variable	Single rater RC, degrees			Intra-rater RC for all raters, degrees		
	Rater 1	Rater 2	Rater 3	Pre-OP	Post-OP	Overall
T1–T12	5	11	9	9	9	9
T2–T12	6	12	13	12	9	11
T4–T12	5	7	10	9	6	8
T5–T12	5	7	10	10	5	8
T1-nonfixed	6	12	10	10	10	11
Nonfixed TK	6	12	13	12	10	11
L1–S1	6	6	7	9	12	11
Nonfixed LL	6	9	14	9	11	10

RC reliability coefficient, TK thoracic kyphosis, LL lumbar lordosis

et al. [35], the authors found an intra-rater agreement of 6° and 4° for T1–T12 and T4–T12, respectively, and an inter-rater agreement of 7° and 6°, respectively. This study, however, utilized EOS-imaging, which is a slot-scanning device that may improve the agreement. Moreover, EOS is currently only available in selective centers. Similarly, Kukko et al. [36] found that the intra-rater agreement for T2–T12 and T5–T12 was 5° and 6°, respectively. However, none of these studies addressed the issue of random effects, so these results are not directly comparable to the present study and likely to underestimate the overall variation seen between randomly chosen raters. Carman et al. [37] measuring nonfixed TK found 95% of the differences between raters to be within 7° and found a trend towards less variation with increased clarity on radiographs. The study also found that an 11° difference in TK was required to rule out measurement error with 95% confidence. Our results are in line with these findings showing that TK measurements have considerable intra- and inter-rater variation and a difference of 8° to 13° (depending on TK definition) may solely be produced by observer error alone.

In order to ensure clinical applicability of our results, our study included both pre- and postoperative radiographs.

**Table 3** Inter-rater RC between three raters with pre- and postoperative subgroups

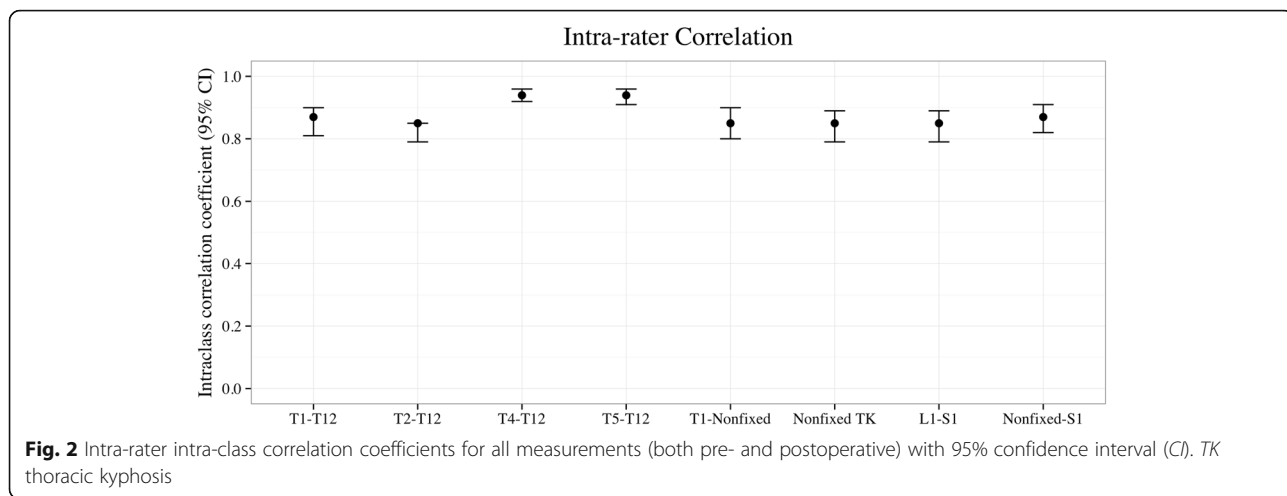
Variable	Inter-rater RC, degrees		
	Pre-OP	Post-OP	Overall
T1–T12	11	14	13
T2–T12	12	10	11
T4–T12	12	7	10
T5–T12	11	7	9
T1-nonfixed	12	14	13
Nonfixed TK	14	13	13
L1–S1	9	12	11
Nonfixed LL	11	12	11

RC reliability coefficient, TK thoracic kyphosis, LL lumbar lordosis

Our analysis showed substantial differences in both intra- and inter-rater agreement between pre- and postoperative radiographs, showing markedly better agreement in postoperative radiographs for T4–T12 and T5–T12 (Tables 2 and 3). For the remaining TK measurements, analyses of pre- and postoperative subgroups were not conclusive but, generally, we found poorer or unchanged agreement. The reason for these changes may be that the variation seen in T4–12 and T5–T12 is mainly due to the lateral and rotational deformity of the curve which is surgically corrected whereas the variation seen in measurements including T1 and T2 is more likely due to structural overlap (e.g., of the humeral head) and therefore not affected by surgery. Interestingly, our analysis also showed considerable variation for the fixed and nonfixed LL, indicating that the sagittal radiograph, as a whole, is inherently difficult to analyze in a reproducible manner in AIS patients.

Our study focused on the overall TK because we found that a wide range of definitions exist in the literature. Establishing the respective reproducibility of these measurements was our main objective, but we would encourage future studies to include additional clinically important parameters, such as proximal TK (T2–T5) and thoracolumbar alignment (T10–L2) as well as several other clinically relevant measurements.

The ICC analysis showed good to excellent reliability for all measurements. However, while the ICC analysis is frequently reported in studies of this type, it holds limited practical value when assessing potential variation of individual measurements per subject, as it is a measure of the reliability for the measurement to differentiate between subjects. By applying a mixed effects model to our data, the observed variance is split into both the variability between the raters within subjects (inter-rater variation) and a residual error term (representing intra-rater variation). Ultimately, an RC is generated which represents the upper and lower 95% limits of agreement for an individual measurement. By using the rater as random effects, our results represent conservative estimates

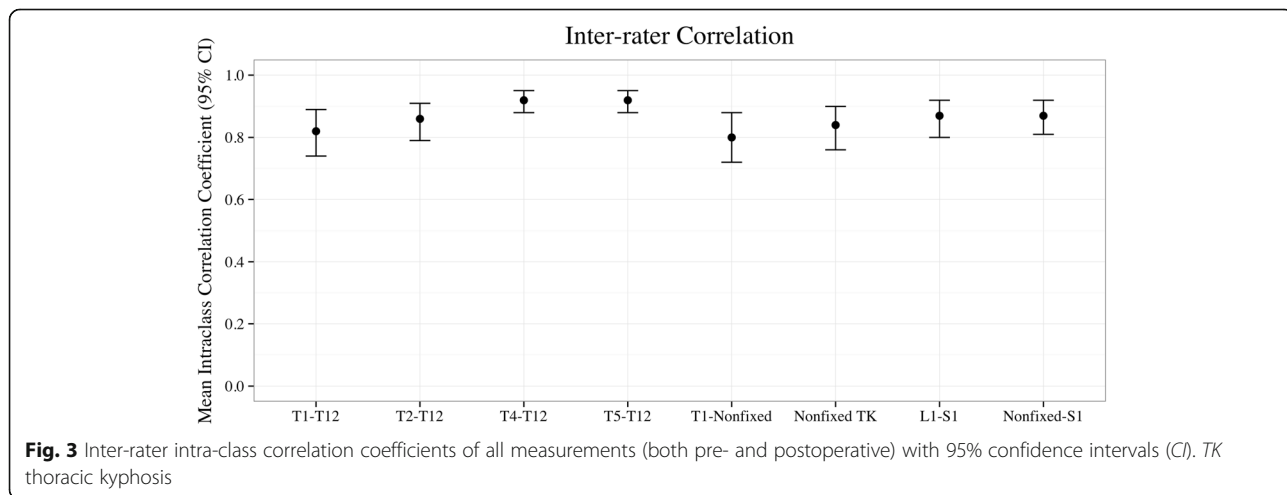


and we hypothesize that measurement variation found in our study would also apply for other raters.

Our results are limited by a substantial variation in single rater RC among raters, which was lowest for T4–T12 and highest for T2–T12 (Table 2). Several steps were taken before the initiation of the study to minimize bias in terms of discrepancies in labeling vertebra, handling of odd number of ribs, or definitions using the nonfixed approach. Rater 1 had more than 3 years of experience in evaluating radiographs from AIS patients, and raters 2 and 3 had 8 and 10 years of experience, respectively. It should be noted that all raters routinely use mainly T5–T12 or T2–12 when evaluating patients with AIS although rater 1 also uses the nonfixed approach on a regular basis. As such, we believe that our results reflect the expected variation between clinicians. In addition, our patient sample did not include lumbar curves, so we cannot infer that our results may be readily applied to this group. Also, the sample size in our study did not allow for analyzing individual curve types, but variation may be greater for

thoracic curves since TK has been found to depend on curve type [38]. Nonetheless, we hope that our study can form the foundation whereby future studies can further elaborate upon different curve types.

Our results may guide clinicians and researchers in the evaluation of the sagittal profile following surgery in defining the limits of actual improvement of worsening of TK as opposed to expected measurement variation. Applying such variation in clinical definitions of progression has previously been described in guidelines for evaluation of radiographic results of brace treatment [37, 39, 40]. Our results indicate that T4–T12 and T5–T12 offer the least amount of observer variation, and while measuring nonfixed TK may offer a more individualized assessment of the spine, we found considerable measurement variation using this approach that may limit the clinical applicability. It is outside the scope of this paper to determine how these various measurements correlate with clinical outcomes; however, we recommend that future studies specifically state the applied definition



of TK and also adequately address measurement variation when evaluating treatment results. This will further help with standardization of measurements between studies for comparative purposes.

## Conclusions

Our study addresses the intra- and inter-rater reproducibility of TK measurements in AIS patients, and we noted a considerable variation for all TK measurements. Both intra- and inter-rater reproducibility were best for T4–T12 and T5–12. Future studies should consider adopting a relevant minimum difference (depending on TK definition) as a limit for indication of true change in TK within a patient. As such, our findings have implications in the decision-making of the spine specialist.

## Additional file

**Additional file 1:** Original primary data. (XLSX 29 kb)

## Abbreviations

AIS: Adolescent idiopathic scoliosis; CI: Confidence interval; ICC: Intra-class correlation coefficient; LL: Lumbar lordosis; RC: Reliability coefficient; SD: Standard deviation; TK: Thoracic kyphosis

## Acknowledgements

There are no further individuals or entities we wish to acknowledge that have played a role in the development and completion of our work.

## Funding

SO was supported by a grant from the Lundbeck foundation. The following grants were received outside the submitted work: SO: Research grant from K2M; DWH: Research grant from Globus Medical; MG: Grant from Globus Medical, K2M, and Medtronic; BD: Grant from Globus Medical, K2M, and Medtronic.

## Availability of data and materials

All original data used for the analyses presented are made available as Additional file 1.

## Authors' contributions

SO, DS, DH, and JC participated in the conception and design of the study. SO, JC, and KK made radiographic measurements. SO and DH analyzed the data. SO, DS, DH, JC, and KK wrote the manuscript, and BD, MG, and KC made critical revisions. DS supervised the study and provided administrative support. All authors have read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

All authors have reviewed the work and agree with its contents.

## Ethics approval and consent to participate

The study was approved by the local institutional review board.

## Disclosures

Søren Ohrt-Nissen was supported by a grant from the Lundbeck foundation and a research fellowship from K2M.

Received: 6 June 2016 Accepted: 14 February 2017

Published online: 21 February 2017

## References

1. Perdriolle R, Vidal J. Morphology of scoliosis: three-dimensional evolution. *Orthopedics*. 1987;10:909–15.
2. Fletcher ND, Jeffrey H, Anna M, et al. Residual thoracic hypokyphosis after posterior spinal fusion and instrumentation in adolescent idiopathic scoliosis. *Spine (Phila Pa 1976)*. 2012;37:200–6.
3. Hwang SW, Samdani AF, Marks M, et al. Five-year clinical and radiographic outcomes using pedicle screw only constructs in the treatment of adolescent idiopathic scoliosis. *Eur Spine J*. 2013;22:1292–9.
4. Lykissas MG, Jain VV, Nathan ST, et al. Mid- to long-term outcomes in adolescent idiopathic scoliosis after instrumented posterior spinal fusion: a meta-analysis. *Spine (Phila Pa 1976)*. 2013;38:E113–9.
5. Hwang S, Samdani A, Tantsorski M, et al. Cervical sagittal plane decompensation after surgery for adolescent idiopathic scoliosis: an effect imparted by postoperative thoracic hypokyphosis. *J Neurosurg Spine*. 2011;15:491–6.
6. Ilharborde B, Morel E, Mazda K, et al. Adjacent segment disease after instrumented fusion for idiopathic scoliosis: review of current trends and controversies. *J Spinal Disord Tech*. 2009;22:530–9.
7. Cao Y, Xiong W, Li F. Pedicle screw versus hybrid construct instrumentation in adolescent idiopathic scoliosis: meta-analysis of thoracic kyphosis. *Spine (Phila Pa 1976)*. 2014;39:E800–10.
8. Boseker EH, Moe JH, Winter RB, et al. Determination of “normal” thoracic kyphosis: a roentgenographic study of 121 “normal” children. *J Pediatr Orthop*. 2000;20:796–8.
9. Propst-Proctor SL, Bleck EE. Radiographic determination of lordosis and kyphosis in normal and scoliotic children. *J Pediatr Orthop*. 1983;3:344–6.
10. Gelb DE, Lenke LG, Bridwell KH, et al. An analysis of sagittal spinal alignment in 100 asymptomatic middle and older aged volunteers. *Spine (Phila Pa 1976)*. 1995;20:1351–8.
11. Rose PS, Lenke LG. Classification of Operative Adolescent Idiopathic Scoliosis: Treatment Guidelines. *Orthop Clin North Am*. 2007;38:521–9.
12. Dang NR, Moreau MJ, Hill DL, et al. Intra-observer reproducibility and interobserver reliability of the radiographic parameters in the Spinal Deformity Study Group's AIS Radiographic Measurement Manual. *Spine (Phila Pa 1976)*. 2005;30:1064–9.
13. Kuklo T, Potter B, Polly D et al. Reliability analysis for manual adolescent idiopathic scoliosis measurements. *Spine (Phila Pa 1976)*. 2005;30:444–54.
14. Wu W, Liang J, Du Y, et al. Reliability and reproducibility analysis of the Cobb angle and assessing sagittal plane by computer-assisted and manual measurement tools. *BMC Musculoskelet Disord*. 2014;15:33.
15. Basques BA, Long WD, Golinvaux NS, et al. Poor visualization limits diagnosis of proximal junctional kyphosis in adolescent idiopathic scoliosis. *Spine J*. 2015. Epub ahead of print.
16. Cheung J, Wever DJ, Veldhuizen AG, et al. The reliability of quantitative analysis on digital images of the scoliotic spine. *Eur Spine J*. 2002;11:535–42.
17. Harrison DE, Cailliet R, Harrison DD, et al. Reliability of centroid, Cobb, and Harrison posterior tangent methods: which to choose for analysis of thoracic kyphosis. *Spine (Phila Pa 1976)*. 2001;26:E227–34.
18. Crawford AH, Lykissas MG, Gao X, et al. All-pedicle screw versus hybrid instrumentation in adolescent idiopathic scoliosis surgery: a comparative radiographical study with a minimum 2-year follow-up. *Spine (Phila Pa 1976)*. 2013;38:1199–208.
19. Yu C-H, Chen P-Q, Ma S-C, et al. Segmental correction of adolescent idiopathic scoliosis by all-screw fixation method in adolescents and young adults. Minimum 5 years follow-up with SF-36 questionnaire. *Scoliosis*. 2012;7:5.
20. Kadoury S, Cheriet F, Labelle H. Prediction of the T2-T12 kyphosis in adolescent idiopathic scoliosis using a multivariate regression model. *Stud Health Technol Inform*. 2008;140:269–72.
21. Voutsinas SA, MacEwen GD. Sagittal profiles of the spine. *Clin Orthop Relat Res*. 1986;210:235–42.
22. Mac-Thiong J-M, Pinel-Giroux F-M, de Guise JA, et al. Comparison between constrained and non-constrained Cobb techniques for the assessment of thoracic kyphosis and lumbar lordosis. *Eur Spine J*. 2007;16:1325–31.
23. Vaz G, Roussouly P, Berthonnaud E, et al. Sagittal morphology and equilibrium of pelvis and spine. *Eur Spine J*. 2002;11:80–7.

24. Dimar J, Carreon L, Labelle H, et al. Intra- and inter-observer reliability of determining radiographic sagittal parameters of the spine and pelvis using a manual and a computer-assisted methods. *Eur Spine J*. 2008;17:1373–9.
25. Cheung KMC, Fracs M, Fhkam F, et al. Predictability of the fulcrum bending radiograph in scoliosis correction with alternate-level pedicle screw fixation. *J Bone Joint Surg Am*. 2010;92:169–76.
26. Samartzis D, Leung Y, Shigematsu H, et al. Selection of fusion levels using the fulcrum bending radiograph for the management of adolescent idiopathic scoliosis patients with alternate level pedicle screw strategy: clinical decision-making and outcomes. *PLoS One*. 2015;10:e0120302.
27. Lenke LG, Betz RR, Harms J, et al. Adolescent idiopathic scoliosis: a new classification to determine extent of spinal arthrodesis. *J Bone Joint Surg Am*. 2001;83:1169–81.
28. O'Brien M, Kuklo T, Blanke T et al. Radiographic measurement manual. Spine Deformity Study Group. Minnesota: Medtronic Sofamor Danek USA Inc.; 2008.
29. Schwab F, Ungar B, Blondel B, et al. Scoliosis Research Society—Schwab adult spinal deformity classification. *Spine (Phila Pa 1976)*. 2012;37:1077–82.
30. Kottner J, Audige L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Int J Nurs Stud*. 2011;48:661–71.
31. Horton WC, Brown CW, Bridwell KH, et al. Is there an optimal patient stance for obtaining a lateral 36" radiograph? A critical comparison of three techniques. *Spine (Phila Pa 1976)*. 2005;30:427–33.
32. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8:135–60.
33. Munro B. Statistical methods for health care research. 3rd ed. Philadelphia: Lippincott-Raven; 1997. p. 224–45.
34. Kuklo TR, Potter BK, O'Brien MF, et al. Reliability analysis for digital adolescent idiopathic scoliosis measurements. *J Spinal Disord Tech*. 2005;18:152–9.
35. Ilharreborde B, Steffen JS, Nectoux E, et al. Angle measurement reproducibility using EOS three-dimensional reconstructions in adolescent idiopathic scoliosis treated by posterior instrumentation. *Spine (Phila Pa 1976)*. 2011;36:E1306–13.
36. Kuklo TR, Potter BK, Schroeder TM, et al. Comparison of manual and digital measurements in adolescent idiopathic scoliosis. *Spine (Phila Pa 1976)*. 2006;31:1240–6.
37. Carman DL, Browne RH, Birch JG. Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobserver variation. *J Bone Joint Surg Am*. 1990;72:328–33.
38. Charlebois M, Mac-Thiong J-M, Huot M-P, et al. Relation between the pelvis and the sagittal profile in adolescent idiopathic scoliosis: the influence of curve type. *Stud Health Technol Inform*. 2002;91:140–3.
39. Richards BS, Bernstein RM, D'Amato CR, et al. Standardization of criteria for adolescent idiopathic scoliosis brace studies: SRS committee on bracing and nonoperative management. *Spine (Phila Pa 1976)*. 2005;30:2068–75. discussion 2076–7.
40. Negrini S, Hresko TM, O'Brien JP, et al. Recommendations for research studies on treatment of idiopathic scoliosis: consensus 2014 between SOSORT and SRS non-operative management committee. *Scoliosis*. 2015;10:1–12.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

