Intervariability in radiographic parameters and general evaluation of a low-dose fluoroscopic technique in patients with idiopathic scoliosis

Christian Wong[1] ⓘ, Jens Adriansen[2], Jytte Jeppsen[3] and Andreas Balslev-Clausen[1]

Abstract

Background: Radiographic images in adolescent idiopathic scoliosis (AIS) have a potential radiation-induced oncogenic effect; thus lowering radiation dose by using fluoroscopic imaging technique of low-dose fluoroscopic technique (LFT) which might be relevant for clinical evaluation.

Purpose: To compare radiographs of LFT with gold standard radiographs for AIS ordinary radiographic technique (ORT).

Material and Methods: Image quality was evaluated for LTF and ORT of a child phantom and two 3D-printed models (3DPSs) of AIS. We measured the primary physical characteristics of noise, contrast, spatial resolution, signal-to-noise ratio, and contrast-to-noise ratio. Three independent evaluators assessed the radiographs by observer-based methods of image criteria (ICS) and visual grading analysis(VGAS). Radiation doses were evaluated by the dose-area-product (DAP) of the 25 phantom radiographs. Reliability and agreement of Cobb's angle (CA) and other radiographic parameters were evaluated on the 3DPSs and reliability on 342 LFT.

Results: The average noise and contrast were approximately 15-fold higher for LFT. SNR and CNR were similar. Overall, ICS and VGAS were 3-fold higher for ORT than for LFT for L3 and similar for Th6. Reliability and agreement were good for the experimental LFT, and the interclass correlation coefficient for CA was 0.852 for the clinical LFT. The average DAP and effective dose for LFT were 8-fold lower than those for ORT.

Conclusion: In conclusion, LFT is reliable for CA measurements and is thus useful for clinical outpatient follow-up evaluation. Even though the image quality is lower for LFT than ORT, the merits are the substantially reduced radiation and a lowered malignancy risk without compromising the measurement of Cobb's angle, thus following the principles of ALARA.

Keywords
Fluoroscopy, spine, technology assessments, radiation safety, dosimetry, skeletal-axial

Introduction

Adolescent idiopathic scoliosis (AIS) is the most commonly occurring structural scoliosis disorder in children and adolescents in pediatric and spinal orthopedic clinics.[1] Patients are often diagnosed in early adolescence with typically smaller curves, where serial screening with repeated spinal radiographs is the only "intervention." This is performed approximately every half year until skeletal maturity if no progression occurs.[1] Initially, postero-anterior (PA) frontal projection radiographs were used for diagnosing AIS as well as identifying the type of curve(-s), determining the severity and identifying other pathologies. Subsequently, serial PA radiographs were used to monitor for potential progression. Cobb's angle (CA) is usually measured to evaluate curve severity.[2]

Increased awareness of the potential radiation-induced oncogenic effect induced by serial radiographs has been raised due to higher incidence rates of leukemia, stomach and upper gastrointestinal cancers, lung cancer, and especially breast cancer[3–6] since the prevalence of AIS in girls is higher than that in boys. This has spawned technical improvements to reduce radiation exposure while optimizing radiographic quality by using PA projection, air-gap techniques, and optimized digital systems; optimizing protocols for the imaging systems and consensus strategies for monitoring scoliosis; and utilizing new methods of recording as the (low dose 2D/3D imaging system EOS system).[7,8] The digital PA ordinary radiographic technique (ORT) is considered the gold standard today and is generally utilized for AIS monitoring. In this study, we will evaluate a low-dose fluoroscopic technique (LFT) using digital screen films. We have utilized LFT for diagnosing gastrointestinal diseases in our clinic, and it is readily available in most hospitals as opposed to, for example, the newer, rarely assessable, and costly EOS technique.[7] Three previous studies in English were conducted using a somewhat similar technique with a focus on image quality and radiation dose but did not utilize LFT in the clinical setting of a longitudinal study.[9–11]

In general, when introducing a radiological method in a new field or optimizing the radiographic quality of an existing method, the image quality must satisfy the diagnostic requirements by retrieving the relevant clinical information while exposing the patient to the lowest possible dose of radiation[12–13]; an image with good quality is an image that fulfills its "diagnostic purpose." This also means that a noisy image might fulfill the diagnostic purpose while not being an image with the highest possible spatial resolution and lowest noise.[14] This is formulated in the radiographic principle of ALARA (as low as reasonably achievable).[15] In AIS, the initial purpose is to determine whether AIS is present and subsequently to determine curve type and severity as well as other pathologies; this would require a high spatial resolution and low noise to distinguish the substructures of the spine due to the failure of formation or segmentation or to identify bony vertebral tumors or other pathologies. For the subsequent clinical follow-up, the aim is to detect the progression of AIS by measuring CA;

thus, the prerequisite for image quality would seem lower,[12–13] namely, to be able to visualize the endplates (or the pedicles) for measurement of the largest attainable and measurable angle(-s) of the spinal deformity(-ites).[2] Ideally, patients should then be exposed to the least attainable radiation dose to fulfill the abovementioned diagnostic purposes while minimizing the risk of secondary malignancy. This group of adolescent females was exposed to repetitive radiographs of the trunk at a young age.[1,3–6] This justifies the use of a lower-dose technique if sufficient or adequate image quality can be obtained. The present study aims to evaluate whether a LFT is adequate as a clinical radiographic for initial evaluation and/or the subsequent follow-up monitoring of AIS by evaluating image quality, reliability, and agreement and measuring radiation dose.

Materials and methods

Study design

This study was divided into two parts: an experimental part and a clinical part. The aim of the experimental part was to evaluate the physical characteristics of image quality and observer-based evaluation of the LFT and ORT using a pediatric trunk phantom and to measure the radiation exposure. Image quality was also evaluated for radiographs of two "3D-printed" models of scoliotic spines (3DPSs), where the agreement and reliability for LFT and ORT were determined. The aim of the clinical part was to evaluate the inter-rater reliability of the actual clinical examinations for LFT radiographs of AIS from our outpatient clinic. Ideally, "agreement" should be determined on the clinical radiograph using both the LFT and ORT techniques. However, this would require "simultaneous" double examinations, thus exposing the participants to excess radiation. This was deemed ethically unviable; instead, we performed double examinations on the 3DPS. Reliability and agreement were evaluated as proposed by Langensiepen et al.[2]

This study was evaluated and approved by the local Research Ethics Committee (Journal number: 17025334).

Imaging system setups of the LFT and the ORT

The examinations were performed in either the posteroanterior or anteroposterior projection for the ORT and LFT, respectively. The patient was standing facing the image intensifier/detector or radiation tube with extended hips and knees and with their feet 10 cm apart. Two lead aprons were placed in the interscapular (mammary) and sacral (genitals) regions. The distances to the radiation tube (source to detector) were 100 cm for LFT and 230 cm for ORT. For the LFT, two images of the thoracic and lumbar spine were necessary to cover the thoracolumbar spine. For ORT, only one exposure was needed. The LFT was performed on a

[1]Department of Orthopedics, University Hospital of Hvidovre, Hvidovre, Denmark
[2]Department of Radiology, University Hospital of Hvidovre, Hvidovre, Denmark
[3]Center for Health Technology, University Hospital of Hvidovre, Hvidovre, Denmark

Corresponding author:
Christian Wong, Department of Orthopedics, University Hospital of Hvidovre, Kettegaard Allé 30, Hvidovre 2650, Denmark.
Email: cwon0002@regionh.dk

DelftDI D2RS system with fluoroscopic exposure of 4 pulses per second/frame per second, and the ORT was performed using a digital Carestream DRX-Evolution system with automatic exposure control. The following acquisition parameters were assessed: tube potential (85 kVp, density set at low for LFT and 71 kVp, density set at 0 for ORT), both with grid, using manual exposures (LFT) and automated mA selection exposure control (ORT) and no additional filtration for LFT and additional filtration (1 mm Al + 0.1 mm Cu) for ORT. The pixel size of 0.139 mm for ORT gives a maximum resolution of 3.6 lp/mm. The pixel size of 0.159 mm for LFT gives a maximum resolution of 3.1 lp/mm.

Evaluators

In the experimental part, observer-based assessments were performed once by 3 evaluators. Evaluator 1 was an experienced (22-year practice) pediatric orthopedic surgeon, evaluator 2 was an experienced (16-year practice) pediatric orthopedic surgeon, and evaluator 3 was a reporting radiographer with 10 years of practice. Evaluator 1 conducted all analyses for the primary physical characteristics. For the 3DPS, measurements of CA and classification according to Nash and Moe (NM) were conducted three times as single sessions at least 1 week apart. In the clinical part, the clinical radiographs were analyzed in a single session by evaluators 1 and 2. Evaluators were blinded to patient identity and clinical information. All image evaluations were performed on a PACS system (Impax 6.4.0, Agfa ® HealthCare, Mortsel, Belgium) on a 3-megapixel viewing station by the 3 evaluators separately.

Part 1. Evaluation of image quality of the LFT and ORT imaging systems

The pediatric trunk phantom. The Pediatric Whole Body Phantom "PBU-70" was examined radiographically for image quality. We chose only the phantom torso of a 4-year-old child with a height of 105 cm with life-size, full-body anthropomorphic measurements with embedded soft tissue substitutes of a synthetic skeleton, lungs, liver, mediastinum, and kidneys (phantom).[16] We performed 25 radiographic examinations with both techniques, where the phantom was placed and repositioned for every recording.

The two 3D-printed AIS models. Radiographs of the two 3DPSs with small and severe lumbar scoliosis were recorded using both techniques. We performed 1 set of radiographs with both techniques. The three evaluators performed 3 measurements of CA and NM at least 1 week apart. Inter- and intra-rater reliability for the three evaluators were assessed using analysis of interclass correlation as well

as the mean absolute difference (MAD), standard error of measurement (SEM), and Bland–Altman plots for the mean differences with additional analysis for systematic differences. We defined accuracy from direct measurement with a protractor for medical purposes. This was seen as a surrogate measure of overall accuracy. The phantom and 3DPSs are shown in Figure 1, and the radiographs are shown in Figure 2.

Primary physical characteristics of image quality evaluation: noise, contrast, and SNR. We examined the imaging characteristics of the LFT and ORT radiographs of the phantom and the two 3DPSs. Twenty-five consecutive radiographs of the phantom using both techniques were evaluated. The objective primary physical characteristics were the contrast, random noise, signal-to-noise ratio, and contrast-to-noise ratio. The contrast was defined as the relative signal difference between the two predefined locations of the bony vertebral spine and the surrounding adjacent tissue of the spine, namely, a square region of interest of the vertebral spine, where the upper and lower pedicles and endplates were included, along with a similar region of interest of the adjacent soft tissue, where no bony soft tissue was included (see Figure 2). Random noise was defined as the fluctuations of the signal over the image when uniformly exposed, as expressed by the standard deviation.[14,17]

SNR was defined as the ratio of the signal (defined as the signal of the object) divided by the standard deviation (of the background).[17] CNR was defined as the ratio of the signal (defined as the difference of the signal of the object and the background) divided by the standard deviation (of the background). Figure 2 shows these regions of interest and radiographs of the phantom, the 3DPSs, and the clinical radiographs using the LFT.

Observer-based methods of image quality evaluation using the visibility of anatomical structures. For ethical reasons, we only performed observer-based evaluations on the images from the phantom and the 3DPSs since we did not want to expose patients to double examinations of the ORT and LFT. We utilized two observer-based methods for comparison of image quality, namely, scores of the image criteria (ICS) and visual grading analysis (VGAS).[14,18] We compared a reference LFT image to all 25 ORT images as well as an ORT image to all 25 LFS images. The reference images were selected randomly.[14]

Using the ICS, the absolute level of image quality of specific structures was determined by evaluating a specific structure compared to the same structure in the reference image by the observer; thus, the observer's decision threshold was constant. The task of the observer was to decide whether the specific structure was superiorly visually reproduced[1] or inferiorly visually reproduced when compared to the reference (0) (the criterion). We
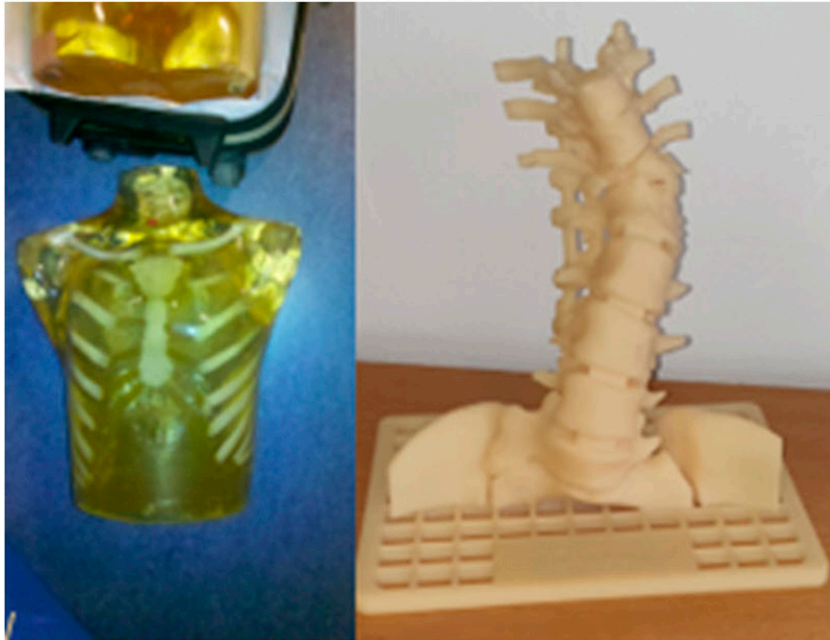
Figure 1. The pediatric whole-body phantom "PBU-70" and one of the two 3D-printed scoliosis models.

Figure 2. Radiographs using the LFT (darker images) and the ORT (brighter images) of the 3D-printed scoliosis models (top four left) and the phantom (top two right). Radiographs (ORT) of scoliosis with regions of interest (below left) and clinical LFT radiographs (below right).

calculated the ICS for the lumbar (L3) and thoracic (Th6) vertebrae as

$$ICS = \frac{\sum_{i=1}^{I} \sum_{c=1}^{C} \sum_{o=1}^{O} F_{i,c,o}}{I \times C \times O}$$

where $F_{i,c,o}$ = fulfillment of criterion c for image i and observer o. $F_{i,c,o} = 1$ if criterion c is fulfilled; otherwise, $F_{i,c,o} = 0$, I = number of images, C = number of criteria, and O = number of observers. The number of images assessed for both the LFT and ORT was 25 ($I = 25$), the number of evaluators was 3 ($O = 3$), and the numbers of criteria were 6 for the thoracic spine and 7 for the lumbar spine ($C = 6$ for the thoracic spine and $C = 7$ for the lumbar spine) in accordance with the chosen CEC criteria.

We also graded the appearance using the VGA method by grading the visibility of specific anatomical structures between images of the two radiographic techniques according to Månsson.[18] The visibility was graded according to a five-level scale: clearly inferior visibility of a specific structure in the image compared to the same structure in the reference image (−2), slightly inferior to (−1), equal to (0), slightly better than (+1), or clearly better than (+2). We calculated the VGAS for the lumbar (L3) and thoracic (Th6) vertebrae as

$$VGAS = \frac{\sum_{i=1}^{I} \sum_{s=1}^{S} \sum_{o=1}^{O} G_{i,s,o}}{I \times S \times O}$$

where $G_{i,s,o}$ = grading (−2, −1, 0, +1, or +2) for image i, structure s, and observer o; I = number of images; S = number of structures; and O = number of observers. The number of images assessed for both the LFT and ORT was 25 ($I = 25$), the number of evaluators was 3 ($O = 3$), and the numbers of criteria were 6 for the thoracic spine and 7 for the lumbar spine ($C = 6$ for thoracic spine and $C = 7$ for lumbar spine) in accordance with the chosen CEC criteria. The specific evaluated (vertebral) structures were in accordance with the European guidelines on Quality Criteria for Diagnostic Radiographic Images in 1990 and 1996.[19–21] The *CEC lumbar spine criteria, 1990,* were chosen for the thoracic evaluation of the images since they focused on osseous vertebral spine structures as well as adjacent soft tissues. For the lumbar evaluation, the *CEC lumbar spine criteria, 1996,* were chosen since they also included an evaluation of the sacroiliac joints.

Radiation dose. The recording systems of LFT and ORT measured the dose area product (DAP) by an integrated DAP meter, and these were stored with the images. Monte Carlo calculations using X-ray dosimetry software (PCXMC, version 2.0; Stuk, Helsinki, Finland) were utilized to determine the effective radiation doses. This was based on the recorded DAP and the principal patient size of the phantom as well as technical and geometric exposure parameters.

Part 2: Evaluation of AIS radiographs using the LFT for clinical reliability

Clinical routine radiographs using LFT. Low-dose fluoroscopic technique was our regular routine clinical radiographic method for AIS for 3 years. One hundred thirty-six adolescent patients with AIS were included as participants. The sex ratio F:M was approximately 2:1. The average age was 13.4 years (range 6–17). We retrieved 342 LFT images of 136 patients with AIS. Two independent evaluators performed measurements of six radiographic parameters once and separately, where they were blinded to the clinical data and previous evaluations. The parameters were the CA, the level of the upper and lower vertebrae used for determining CA, the NM, and the Metha angles on the left and right sides at the apex vertebrae. Table 1 shows the radiographic characteristics of the participants. Figure 2 illustrates the standard clinical LFT radiographs.

Statistical analyses. Reliability was assessed using the interclass correlation coefficient (ICC). In part 1 of the two 3DPS models, inter- and intra-reliability for CA and NM were assessed using a 2-way mixed model for consistency for the 3 evaluators for their 3 separate measurements. Agreement was defined as the MAD, SEM, and Bland–Altman plots for the mean differences with additional analyses of one-way t-test and logistic regression for significant and systematic differences, respectively. In part 2 of the clinical evaluation, the single measurements of CA and the other 5 radiographic parameters of the two evaluators

Table 1. Distribution of participants according to the classification of King and Moe. Type 1: an "S"-shaped deformity, in which both curves are structural and cross the CSVL (midline), with the lumbar curve being larger than the thoracic curve. Type 2: an "S"-shaped deformity, in which both curves are structural and cross the CSVL, with the thoracic curve being larger than or equal to the lumbar curve. Type 3: major thoracic curve in which only the thoracic curve is structural and crosses the CSVL. Type 4: long "C"-shaped thoracic curve in which the fifth lumbar vertebra is centered over the sacrum and the fourth lumbar vertebra is tilted into the thoracic curve. Type 5: double thoracic curve.

	Dextro convex	Sinistro convex
Type 1	7	29
Type 2	10	0
Type 3	24	5
Type 4	22	28
Type 5	2	1
Unclassifiable	8	

were assessed using a 2-way mixed model of ICC for absolute agreement for inter-rater reliability.

Inter- and intra-rater reliability assessed by the ICC was considered with the following limits of agreement: poor; 0.0–0.20 slight; 0.21–0.40 fair; 0.41–0.60 moderate; 0.61–0.80 substantial; and 0.81–1 almost perfect. Independent sample t-tests or Mann–Whitney nonparametric tests were conducted for differences in SNR and CNR, differences in image quality, and radiation dose. This depended on if normal distribution was present. This was tested by Kolmogorov–Smirnov and Shapiro–Wilk tests and evaluation of QQ plots. We considered a p-value of <0.05 as a significant result. ICC and other statistical evaluations were performed using IBM SPSS Statistics, version 22 (IBM©, Chicago, Il, USA).

Results

Part 1: Evaluation of image quality and dose evaluation

Phantom images: Noise, Contrast, and SNR. We performed 25 LFT and 25 ORT radiographs of the PBU-70 phantom for both dose measurements and image quality analyses. Table 2 shows the averages and ranges of noise for the vertebral spine and soft tissue and the calculated SNR for the phantom images.

None of the parameters were normally distributed; thus, Mann–Whitney tests were performed. The average noise levels for LFT were almost 14-fold higher for bone, 19-fold higher for soft tissue, and 13-fold higher for contrast, which were all significantly different (<0.001). The SNR and CNR were approximately similar for both techniques and not significantly different. Figure 3 shows the variation in

contrast for the 25 images for the LFT and ORT, where there is a larger variation for the LFT.

3DPS: Noise and SNR. We performed 1 LFT and ORT of the two 3DPSs for image quality analyses. Table 2 shows the averages and ranges of noise for the vertebral spine and soft tissue and the overall calculated SNR for 25 images for the two 3DPS.

The average noise for bone was markedly higher for both bone and soft tissue for the ORT. The SNR was also markedly higher for ORT.

VGAS and ICS. All LFT and randomly chosen ORT images and vice versa were compared. The overall values for ICA and VGAS as well as for the specific structures (each criterion) are shown in Table 3.

Upon evaluating the 25 images of the phantom, the overall ICS and VGAS (negative values for VGAS) were 3-fold higher for ORT than for LFT for L3. The overall ICS and VGAS were approximately similar for Th6 cells. When looking at the specific structures, there were similar or better evaluations for endplates, pedicles, and lateral cortex in IC and VGA for both Th6 and L3 for the LFT and for the other structures in IC and VGA for both Th6 and L3 for the ORT. When comparing IC and VGA for Th6 and L3, there were significant differences for pedicles, intervertebral joints, and processi spinosus and transversus. There were no significant differences in IC and VGA for endplates. All tests were evaluated with independent sample t-tests.

Part 2: Reliability and agreement

3DPS. The three evaluators assessed CA and NM radiographs of the two 3DPSs using LFT and ORT. Table 4 shows

Table 2. (Top) Average noise for bone and soft tissue, SNR and CNR (N- = noise, LFT bone = for bone with the LFT, ORT bone = for bone with the ORT, LFT st = for soft tissue with the LFT, ORT st = for soft tissue with the ORT, Con LFT = Contract for LFT, Con ORT = contrast for ORT, SNR LFT and SNR ORT = SNR for LFT and ORT, respectively, CNR LFT and CNR ORT = SNR for LFT and ORT, respectively; * indicates significant difference). (Below) Noise for the LFT and ORT for the overall image and regions of interest of bone and soft tissue, separately (AVG = average noise, DEV = standard deviation) for the two 3DPSs.

	N-LFT bone	N-ORT bone	N-LFT st.	N-ORT st.	Con LFT	Con ORT	SNR LFT	SNR ORT	CNR LFT	CNR ORT
Avg.	34349	2394	38710	2066	4362	329	4.42	3.95	0.55	0.53
Std.	7838	689	20388	4293	420	74	0.89	0.35	0.64	0.11
Sig.Diff	*		*		*					

	N-AVG LFT	N-DEV LFT	N-AVG ORT	N-DEV ORT
3DPS 1 overall	40272.00	9135.00	2526.00	252.70
Roi 1 vertebrae	2256.00	558.00	29721.00	7189.00
Roi 1 'soft tissue'	1918.00	251.00	45016.00	1530.00
SNR	0.04		−60.53	
3DPS 2	34934.00	8762.00	2471.00	185.70
Roi 2 vertebrae	2638.00	210.00	34802.00	9587.00
Roi 2 soft tissue'	2543.00	22.00	47688.00	1204.00
SNR	0.01		−69.39	

the ICC for the CA measurements between evaluators and over time as well as SEM and MAD.

We also achieved ICCs for NM of 0.97 and 0.86 for intra- (over time) and inter-rater (in between evaluators) reliability, respectively.

Inter-rater reliability for the clinical images. We evaluated the 342 LFT radiographs of 136 participants with measurements of 6 parameters. Two LFT radiographs were excluded
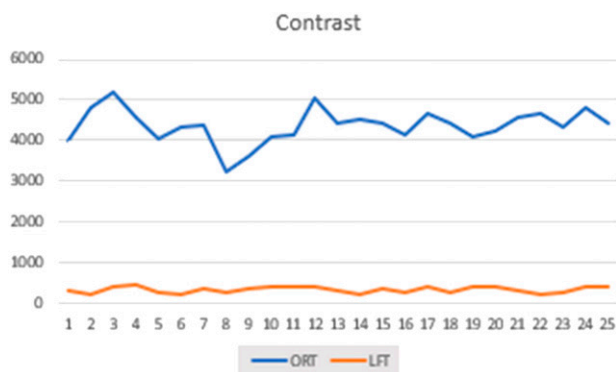

Figure 3. Contrast for the 25 examinations of the phantom using the ORT and LFT.

Table 3. ICS and VGAS for vertebrae L3 and Th 6 for LFT and ORT as well as the ratios for LFT/ORT for ICS and ORT/LFT for VGAS (top) and IC and VGA for the evaluation criteria/specific structures (bottom).

	LFT L3	LFT Th6	ORT L3	ORT Th6
ICS	0.26	0.35	0.74	0.33
VGAS	−0.26	−1	−0.09	1.35
Ratio ICS			2.9	0.94
Ratio VGAS	2.89	0.87		

	LFT-IC	LFT-VGA	ORT-IC	ORT-VGA
Th6				
Endplates	1.04	−0.12	1.04	−0.40
Pedicles	0.20	−0.47	0.08	0.04
Lateral cortex	1.12	−0.96	1.12	−0.20
Interv. Joints	0.16	−0.01	0.16	0.00
Processi T. and S.	0.20	−0.01	0.20	0.00
Adj. Soft Tis.	0.24	−0.01	0.24	0.00
L3				
Endplates	1.04	1.04	1.04	1.00
Pedicles	0.80	1.08	1.08	2.00
Interv. Joints	0.12	0.32	0.32	2.00
Processi T. and S.	0.20	1.16	1.16	1.00
Trab. Struc.	0.20	1.20	1.20	1.04
Adj. Soft Tis.	0.24	0.24	0.24	0.00
sacroilical J.	0.32	1.28	1.28	2.00

due to poor image quality. Additionally, poor quality also led to an inability to evaluate NM and measurements of Metha angles in nine LFT radiographs. Comparisons of measurements for the 6 relevant clinical radiographic parameters for the two evaluators with ICC, MAD, and SEM are shown in Table 4 (top and middle). The average measurement of CA for evaluator 1 was 15.72°, with a standard deviation of 11.63°, and that for evaluator two was 17.05°, with a standard deviation of 12.35°.

Dose measurements. Characteristics for dose measurements for the LFT and ORT are summarized in Table 4. The average DAP for the LFT was 8-fold lower than that for the ORT. The DAP measurements were not normally distributed; thus, the Mann–Whitney test was performed. The DAP doses were significantly different (<0.001). The total average effective dose for LFT was 8-fold lower than that for ORT.

Discussion

In this study, we evaluated the image quality of LFT and ORT imaging systems and the reliability of measurements

Table 4. (Top) ICC of Cobb's angle and classification according to Nash and Moe (Nash and Moes) for the two 3DPS model evaluations (inter = interrater reliability, intra = intrarater reliability, and ACC = accuracy) (Top). ICCs for six radiographic parameters (Nash and Moes = classification according to Nash and Moes, Top vertebrae = the upper vertebrae used for determining CA, Lower vertebrae = the lower vertebrae used for determining CA, Metha (l.) = Metha angle measured on the left side, Metha (r.) = Metha angle measured on the right side, and Diff. in Metha = difference in measured Metha angle) (Middle). Characteristics for dose measurements for the LFT and ORT (Mean DAP in μGy*m², Std. = standard deviation, mean effective dose in mSv) (Below).

	ORT inter	LFT inter	ORT intra	LFT intra	ACC
Cobb's angle	0.97	0.965	0.98	0.982	0.899
Nash and Moes	0.938	1	0.915	0.958	0.968

	ICC
Cobb's angle	0.852
Nash and Moes	0.454
Top vertebrae	0.66
Lower vertebrae	0.806
Metha (l.)	0.454
Metha (r.)	0.156
Diff. in metha	0.814

	DAP (μGy*m2)	Std.Dev	E (mSv)
ORT	9.70	29.9	0.062977
LFT	1.21	7.3	0.007675

of radiographic parameters of AIS in clinical and experimental settings. Our purpose was to examine whether a LFT for AIS (LFT) had the prerequisites to be an adequate clinical radiographic method for evaluating AIS. When we examined image quality, we found that the contrast levels for LFT fluctuated more than those for ORT, where the latter curve was generally more linear. A more varied quality for LFT would indicate either over- or underexposure of the images.[14] The noise levels for bone, soft tissue, and contrast were markedly higher and significantly different for the LFT. The SNR and CNR were similar and not significantly different. Both were not larger than 5, which is the threshold for reliability for human detection,[22] and earlier studies have reported an even higher SNR when used for AIS purposes.[17] The SNR was marginally higher for the LFT, thus suggesting a better ability to reproduce low-contrast objects.[14] This is in accordance with previous studies.[9,10,14] In general, ORT is superior in image quality when compared to LFT based on primary physical characteristics but was approximately similar for key parameters of ratios of signal and contrast to noise.

The observer-based visualizations of overall ICS and VGAS were almost 3-fold higher for the evaluated lumbar vertebrae using the ORT but were almost similar for the evaluated thoracic vertebrae. The ICS for ORT agreed with Tingberg (1994) but was markedly lower for LFT [Figure 4, p 33,14]. This would suggest that ORT is superior in observer-based evaluation for lumbar anatomical structures, indicating that structures are more poorly reproduced and less sharp for LFT, but surprisingly almost equal for thoracic anatomical structures (∼0.9), especially since overall noise levels for bone and soft tissue were higher for LFT in both experimental settings (phantom and 3DPS). Moreover, when evaluating the specific structures with IC and VGA, the LFT was similar or better for determining the endplates, lateral cortex, and pedicles. LFT was significantly different for the endplates. This would indicate that LFT is good for specific purposes, whereas ORT is good for others. We conducted double comparisons of ORT compared to LFT and vice versa with ICS and VGAS. In these double comparisons, our findings were consistent, thus strengthening our findings for ICS and VGAS.

In this study, our mean DAP for ORT was within the range of similar digital radiographic examinations of AIS from previous studies.[17,23] For LFT, our mean DAP was higher than that in studies using microdose EOS or similar LFT protocols,[9–19,24,25] indicating that further optimization could be obtained. However, we found an 8-fold and significantly different decrease in radiation dose in LFT when compared to ORT for mean effective doses. This had consequences for image quality, as described earlier, which were reflected in our clinical evaluation since the reliability for the clinically relevant radiographic parameters was only fair, except for Cobb's angle. This suggests that ORT in the clinical setting is superior to FLT for the initial radiographic examination for excluding other pathologies as failure of formation or segmentation as well as if the intended use is aimed at evaluating radiographic parameters other than CA, that is, for research purposes or surgical evaluation. For these purposes, ORT should still be the method of choice. *We suggest that the LFT is adequate for monitoring since the primary purpose is to monitor the progression of curve severity* by longitudinal measurements of CA; we had "almost perfect" reliability with an ICC of 0.85 for CA in the evaluation of the clinical images. Especially when evaluated from the perspective of ALARA, LFT would seem to have adequate reliability for CA measurements while having a markedly lower radiation dose and hence a lower malignancy risk. We were unable to perform further calculations of effective dose and the Monte Carlo calculations for the clinical radiographs (even though it has a remarkably good correlation with the effective dose) since we did not have the anthropomorphic measures of our patients.[9]

In this study, we also found that the LFT technique produces a short series of sequential radiographs automatically, thus providing the possibility of dynamic radiographic evaluation of the spine. This would seem interesting for research purposes when the dynamic movements of the spine could be evaluated.

The overall ICC for inter- and intra-rater reliability for CA and NM was "almost perfect" for the 3DPS measurements. Surprisingly, the MAD and SEM for ORT were higher than those for LFT. In agreement in the experimental assessments, the three evaluators demonstrated an MAD in the range of that in previous studies, with an MAD between approximately 2.2–5.1° for LFT,[2] whereas the MAD for ORT was generally higher. The SEM values for both LFT and ORT were within the range of previous studies (4.13–5.08).[2] For the clinical images using LFT in clinical patients, almost all evaluated radiographic parameters were superior for ORT, except for CA. Here, we demonstrated an "almost perfect" clinical reliability with an ICC of 0.85, which is considered acceptable reliability and in the range of previous studies.[2] However, there was only "moderate" reliability for the other parameters of the classification of NM and the Metha angles. We considered the high ICC for the difference in Metha angle as a surrogate finding (but clinically important parameter) since it was derived from two "inadequate" measurements of the left and right Metha angles. For agreement in the clinical radiographs, the evaluations of MAD and SEM were within the range of those in previous studies.[2] The latter would indicate some consistency of measurements between the 2 evaluators despite a high SD of the two evaluators of approximately 10°. There were no systematic differences in the Bland–Altman plots and subsequent linear regression analyses between the LFT and ORT methods (see Supplemental Material). In conclusion, the measurements of CA are

reliable in both the experimental and clinical settings for LFT but unreliable for other clinically relevant radiographic parameters.

The LFT has obvious potential causes for poor reliability, namely, technical, that is, the distance of the X-ray tube from the body, the timing of the radiographs, the routine and experience of the technician and the position of patients,[10,12,13,17] and as eluded earlier, future studies should focus on optimizing our protocol, which would potentially reduce the radiation dose further. In this study, we tried to compensate for technical bias by using specially trained radiographic technicians for the clinical radiographs and one technician for the 3DPS and phantom recordings. In the clinical routine, it was not possible to have the same technicians for all AIS radiographs with LFT, and we ascribe the excluded radiographs and inability to measure the Metha angles to this limitation. The LFT technique is dependent on the experience of the technician since the recordings are based on manual exposures, thus potentially involving excess radiation doses due to longer exposures. In our setup, we also needed to perform two separate radiographs of the thoracic and lumbar regions with the LFT, which could cause the participants to rotate or move in between the two recordings. However, a previous study has shown that minor rotation is inconsequential for measurements of CA.[26] In our study, we tried to accommodate this by evaluating the two radiographs separately. However, it is possible to "stitch" the two radiographs together, which is useful in routine clinical evaluation, but to evaluate the LFT, we found it more prudent to evaluate them separately. In general, we tried to adhere to the "quality assessment tool for the diagnostic accuracy studies" checklist (QUA-DAS).[27] We considered the LFT as the index test and compared this to the gold standard reference test of ORT. We chose representative evaluators with varied experience levels who were blinded to their prior findings and the findings of other observers. Some of the examinations were randomized, and the methods and results of all assessments were described. Overall, we also achieved acceptable reliability for CA despite these methodological shortcomings.

In this study, we chose to use the *CEC image criteria of the lumbar spine, 1996,* for the lumbar spine since we were able to evaluate more lumbar structures as the Sacro-iliac joints, and the *CEC image criteria for the lumbar spine, 1990,* for our thoracic spine evaluation since we wanted to evaluate bony spinal structures, and not the soft tissue of the lungs and heart. However, we could have chosen to use the revised criteria as proposed by Lanhede et al. (2002) and Tingberg et al.,[19,28] but when evaluating them, we did not find them more specific for scoliosis evaluation purposes. Instead, we chose to evaluate IC and VGA specifically for isolated criteria, that is, endplate visibility and pedicle visibility. We found that the LFT was similar to or better than the ORT for evaluating structures important for the measurement of Cobb's angle. However, our "unconventional" evaluation, as well as our use of different evaluation criteria, could explain the difference in our observer-based evaluations of the thoracic and lumbar spine. Using observer-based evaluations has inherent bias, which depends on the imaging system itself as well as the skill sets and training of the evaluators. For representativeness, we chose evaluators with different skill sets and levels of experience. However, this might also have influenced our observer-based evaluations. Finally, our consistent double examinations raised our confidence in findings of differences in the lumbar region but not in the thoracic region.

In conclusion, we examined an imaging system using the low-dose fluoroscopy technique for radiographic examination of AIS. In conclusion, the LFT is reliable for CA measurement and is thus useful for follow-up evaluation of scoliosis progression in a clinical setting but is not adequate for appreciating the details and pathology of the spinal skeletal structures. Even though the image quality is lower for LFT than ORT, the usefulness in clinical follow-up evaluations is motivated by the lower exposure of radiation doses to adolescent girls with AIS, thus lowering malignancy risk and remaining in accordance with the principles of ALARA.

Declaration of conflicting interests

Funding

ORCID iD

Christian Wong https://orcid.org/0000-0002-3201-5827

Supplemental material

Supplemental material for this article is available online.

References

1. Negrini S, Donzelli S, Aulisa AG, et al. 2016 SOSORT guidelines: orthopaedic and rehabilitation treatment of idiopathic scoliosis during growth. Scoliosis Spinal Disord 2018; 13: 3.

2. Langensiepen S, Semler O, Sobottke R, et al. Measuring procedures to determine the Cobb angle in idiopathic scoliosis: a systematic review. Eur Spine J 2013; 22: 2360–2371.

3. Simony A, Hansen EJ, Christensen SB, et al. Incidence of cancer in adolescent idiopathic scoliosis patients treated 25 years previously. Eur Spine J 2016; 25: 3366–3370.

4. Ronckers CM, Land CE, Miller JS, et al. Cancer mortality among women frequently exposed to radiographic examinations for spinal disorders. Radiat Res 2010; 174: 83–90.

5. Nash CL, Jr, Gregg EC, Brown RH, et al. Risks of exposure to x-rays in patients undergoing long-term treatment for scoliosis. J Bone Joint Surg Am 1979; 61: 371–374.

6. Doody MM, Lonstein JE, Stovall M, et al. Breast cancer mortality after diagnostic radiography: findings from the U.S. scoliosis cohort study. Spine 2000; 25: 2052–2063.

7. Law M, Ma WK, Chan E, et al. Evaluation of cumulative effective dose and cancer risk from repetitive full spine imaging using EOS system: impact to adolescent patients of different populations. Eur J Radiol 2017; 96: 1–5.

8. Knott P, Pappo E, Cameron M, et al. SOSORT 2012 consensus paper: reducing x-ray exposure in pediatric patients with scoliosis. Scoliosis 2014; 9: 4.

9. Geijer H, Beckman K, Jonsson B, et al. Digital radiography of scoliosis with a scanning method: initial evaluation. Radiology 2001; 218: 402–410.

10. Geijer H, Verdonck B, Beckman KW, et al. Digital radiography of scoliosis with a scanning method: radiation dose optimization. Eur Radiol 2003; 13: 543–551.

11. Kluba T, Schäfer J, Hahnfeldt T, et al. Prospective randomized comparison of radiation exposure from full spine radiographs obtained in three different techniques. Eur Spine J 2006; 15: 752–756.

12. Kohn MM, Moores BM, Schibilla H, et al. European guidelines on quality criteria for diagnostic radiographic images in pediatrics. Brussels: European Commission. Retrieved from: https://op.europa.eu/da/publication-detail/-/publication/47eb62b0-698d-4166-bc34-cc3f8d07d2e3 (1990, accessed 28 December 2020).

13. Carmichae JHE, Maccia C, Moores BM, et al. European guidelines on quality criteria for diagnostic radiographic images. Brussels: European Commission. Retrieved from: https://op.europa.eu/da/publication-detail/-/publication/d59ccc60-97ed-4ce8-b396-3d2d42b284be (1996, accessed 28 December 2020).

14. Tingberg A. Quantifying the quality of medical x-ray images–An evaluation based on normal anatomy for lumbar spine and chest radiography. [dissertation]. Malmoo: University of Lund, 2000.

15. Keith KJ, Strauss SC. The ALARA (as low as reasonably achievable) concept in pediatric interventional and fluoroscopic imaging: striving to keep radiation doses as low as possible during fluoroscopy of pediatric patients—a white paper executive summary. Pediatr Radiol 2006; 36: 110–112.

16. Kyoto Kagaku. Pediatric whole body phantom "PBU-70". PH-2C 41350-300. Retrieved from: https://kyotokagaku.com/en/products_data/ph-2c/ (accessed 1 March 2020).

17. Ernst C, Buls N, Laumen A, et al. Lowered dose full-spine radiography in pediatric patients with idiopathic scoliosis. Eur Spine J 2018; 27: 1089–1095.

18. Månsson LG. Evaluation of radiographic procedures. Investigations related to chest imaging. [dissertation]. Göteborg: University of Gothenburg, 1994.

19. Tingberg A, Herrmann C, Lanhede B, et al. Influence of the characteristic curve on the clinical image quality of lumbar spine and chest radiographs. Br J Radiol 2004; 77: 204–215.

20. Maccia C, Nahrstedt U, Moores BM. European commission. CEC quality criteria for diagnostic radiographic images and patient exposure trial. EUR 12952. Brussels: European Commission. Retrieved from: https://op.europa.eu/en/publication-detail/-/publication/2e25c0ba-72cc-40c1-87cc-372a8924aea7 (1990, accessed 1 March 2020).

21. Moores BM, Carmichael JHE, Maccia C. European guidelines on quality criteria for diagnostic radiographic images. Brussels: European Commission. Retrieved from: https://op.europa.eu/da/publication-detail/-/publication/d59ccc60-97ed-4ce8-b396-3d2d42b284be (2000, accessed 1 March 2020).

22. Rose A. The sensitivity performance of the human eye on an absolute scale. J Opt Soc Am 1948; 38: 196–208.

23. Hui SCN, Pialasse JP, Wong JYH, et al. Radiation dose of digital radiography (DR) versus micro-dose x-ray (EOS) on patients with adolescent idiopathic scoliois: 2016 SOSORT_IRSSD "john sevastic award" winner in imaging research. Scoliosis Spinal Disord 2016; 11: 46.

24. Yvert M, Diallo A, Bessou P, et al. Radiography of scoliosis: comparative dose levels and image quality between a dynamic flat-panel detector and a slot-scanning device (EOS system). Diagn Interv Imaging 2015; 96: 1177–1188.

25. Ilharreborde B, Ferrero C, Alison M, et al. EOS microdose protocol for the radiological follow-up of adolescent Idiopathic scoliosis. Eur Spine J 2016; 25: 526–531.

26. Wong C, Hall J, Gosvig K. The effects of rotation on radiological parameters in the spine. Acta Radiologica, 2019 Mar; 60(3): 338–346.

27. Whitling P. Quandas. Bristol. Bristol, UK: University of Bristol. Retrieved from: https://www.bristol.ac.uk/population-health-sciences/projects/quadas/ (accessed 1 February 2021)

28. Almén A, Tingberg A, Besjakov J, et al. The use of reference image criteria in x-ray diagnostics: an application for the optimisation of lumbar spine radiographs. Eur Radiol 2004; 14: 1561–1567.