S. Champain
K. Benchikh
A. Nogier
C. Mazel
J. De. Guise
W. Skalli

Validation of new clinical quantitative analysis software applicable in spine orthopaedic studies

Received: 28 September 2004
Revised: 24 January 2005
Accepted: 28 February 2005
Published online: 17 June 2005
© Springer-Verlag 2005

S. Champain (✉) · K. Benchikh
A. Nogier · W. Skalli (✉)
Laboratoire de Biomécanique,
ENSAM-CNRS, 151, boulevard
de l'Hôpital, 75013 Paris, France
E-mail: champainsabina@hotmail.com
Tel.: +33-1-44-246365
Fax: +33-1-44-246366
E-mail: Wafa.Skalli@paris.ensam.fr

C. Mazel
Institut Mutualiste Montsouris,
Paris, France

J. De. Guise
LIO-ETS Montréal,
Montreal, Canada

Abstract The objective of this study was to evaluate an X-ray films analysis software, i.e. to estimate the reliability and validity of clinical measurements by means of this software. The authors first performed tests of precision and reproducibility of measures. The precision for dynamic modules was estimated at $\pm 2^\circ$ for the lumbar analysis and $\pm 3^\circ$ for the cervical one. Mean reproducibility coefficients calculated for postural modules are about 4° for the angular parameters and 3 mm for the linear ones. We also evaluated clinical applicability of the software through its validity. Reference values calculated on a population of healthy subjects showed agreement with the literature. Then,

when analysing postural X-ray films of severe scoliotic patients, we found that inter-observer reproducibility coefficients show a lower reliability of measurements; the main cause seems to be the low visibility of anatomic landmarks due to the quality of X-ray films and to the degree of deformity. This study allowed to better estimate the reliability and the usefulness of this tool, allowing for multicentric studies and exchanges.

Keywords Spine · X-ray · Quantitative analysis · Software · Motion · Balance

Introduction

The treatment of most spinal diseases is a difficult task because of the rate of immediate, mid- and long-term complications and failures [16]. Surgical planning and clinical follow up often rely on postural and/or dynamic X-ray exams, among other imaging options.

Many authors investigated the radiological parameters liable to rule patient's evolution, such as postural and balance parameters [2, 4, 8, 9, 11, 12, 26]. These parameters might be classified in pelvic parameters, either morphologic (incidence) or positional (sacral slope [4, 11, 12], pelvic angle [8, 9]), and spinal parameters (such as lordosis, kyphosis [4, 11, 12] or plumb lines [8, 9]). Their measurement is tedious and time consuming while the accuracy of the results may be highly variable depending on the method employed. Furthermore, when

trying to compare the results between studies, a major difficulty comes from the differences in the definition, identification and measurement of parameters.

Recent computer-based tools allowed for a rapid progress in quantitative measurements of a large range of parameters [9, 17, 19, 20, 23, 26]. Moreover, these tools allow easy standardisation of X-ray films measurements, which is of paramount importance when comparing studies in multicentric analysis. However, validation of these tools is essential [3, 17, 18].

In 1999, the team of LBM and LIO developed an X-ray films analysis software (SpineView 1.0) allowing for the measurement of the main lumbar dynamic and sagittal postural parameters of the spine. The evaluation of this first software was achieved by comparison with direct measurements on plain X-ray films [17, 18], and showed good reliability of the measurements.

A second version of this software recently has been developed, increasing the number of parameters measured in most clinical situations, i.e. frontal parameters for the study of spinal deformities, cervical and lumbar dynamic study and sagittal postural analysis. Besides, the development focused on the optimization of the landmarks identification, i.e. automatic and semi-automatic edge detection, as well as on the calculation time.

The purpose of the present study was to evaluate Spineview 2.1 software. The main steps were:

- Evaluation of the reliability (precision, reproducibility) [1, 3, 7, 14]
- Evaluation of the validity, by comparing asymptomatic subjects' reference values for main pelvic and spinal parameters to those published in the literature
- Investigation on the limits of its use.

Materials and methods

Presentation of the main features of the SpineView 2.0 software

The software is based on the quantification of vertebral body position and calculation of parameters on five types of X-ray films: cervical and lumbar dynamic (flexion–extension), frontal and sagittal postural X-ray films and also dynamic lateral bending films.

Each analysis takes four steps: anatomic landmarks identification on a digital X-ray film, automatic edge detection of the vertebra [10], validation of the edge detection after visual control and manual correction, then calculation of parameters with regard to the identified contours, with the possibility to record the analysis and the numerical results.

762 parameters are calculated, such as:

1. Angular parameters: intervertebral angles, global spine inclination [23], sagittal curvatures, sagittal tilt [11], frontal curvatures (Cobb angles), vertebral wedging, pelvic postural and morphologic parameters [4, 8, 9, 11, 12].
2. Linear parameters: vertebral and discal heights [15], sagittal vertebral axis (SVA) [5, 8, 9], plumb lines [8, 28], pelvic postural and morphologic parameters [8, 9, 11, 12].
3. Indexes: overhang, listhesis [15], disc/vertebra height ratio.
4. Parameters related to vertebral kinematics on dynamic X-ray films: range of motion (ROM), vertebral translations [15], mean centres of rotation (MCR) [13]. The MCR is calculated only for ROM values superior to 3° [23].

Evaluation of the reliability

Precision tests were performed for dynamic modules using reference measurements on *in vitro* specimens. Reproducibility tests using *in vivo* radiological data were done both for dynamic and postural modules.

Precision tests

A total of 13 anatomic spine segments (8 C1-T2 and 5 T12-sacrum) were harvested, radiographically screened for bony abnormalities, loss of disc height or other signs of abnormal degeneration, and frozen until use. Metallic pins (1 mm diameter) were rigidly fixed into the spinous processes (cervical segments) or into the vertebral bodies (lumbar ones), as shown in Fig. 1. Dynamic (flexion and extension) radiographs were taken on a 0° angle of incidence and focused on C4 and L3 respectively for the cervical and lumbar spine segments [23]. Ranges of motion at each vertebral level (excepting those rigidly fixed to the supports i.e. C2C3 and T1T2 for dynamic cervical radiographs and L5S1 for lumbar dynamic ones) were accurately measured using the metallic pins. The accuracy of this reference measurement was evaluated using reproducibility tests (30 measurements $\times$ 5 levels per measurement = 150 values, 2 SD = $\pm 0.8^\circ$).

The ranges of motion were then measured on scanned X-ray films (Vidar scan, type VXR 12) by means of Spineview 2.1.

Wilcoxon signed-ranks test was performed to assess possible statistical significant difference between the two sets (reference and Spineview) of measurements, and then the standard error of measurement was calculated for each level. In order to assess a global precision estimator for each dynamic module (respectively cervical and lumbar), the ROM's error normal distribution and independence from the intervertebral level was checked (descriptive statistics, Student's *t* test and k-sample analysis of variance).

Finally, the global precision estimator was calculated for a confidence interval of 95% [1, 7, 14].

Reproducibility tests

Random error of measurement was evaluated for each module and for three classes of quality films, representing variable visibility of the vertebrae on the films. The quality classes were: good/routine, poor and very poor, as assessed by a trained spine surgeon. For each module, three single or pairs of X-ray films (respectively for postural and dynamic modules) were randomly chosen (one per quality class). Three observers performed series of 30 repeated analyses on each film of each class, so that for each module we had a total of 3 $\times$ 3 $\times$ 30 i.e. 270 analyses.

Fig. 1 Examples of dynamic X-ray films used for precision tests. **a** Cervical(C1-T2 segments) X-ray films focused on C4. **b** Lumbar (D12-sacrum segments) X-ray films focused on L3

Reproducibility was analysed for each variable, as follows:

1. For the parameters calculated at each vertebral level, we first performed a k-sample comparison of variances, in order to assess whether the variance depends on the level.
2. All variables were submitted to ANOVA, in order to evaluate the intra- and inter-observer agreement. We also calculated the intra-class correlation coefficient ICC [14, 18, 21], which is a complementary estimator for the power of the inter-observer agreement.

Descriptive statistics were used to characterize all the results in terms of mean, standard deviation, SD, coefficient of variation CV (SD/mean) and coefficient of reproducibility for 95% confidence interval, which is equal to the double of the SD (more precisely $1.96 \times \text{SD}$) [1, 3, 7, 14].¹ This coefficient is used as the global reproducibility estimator.

Validity: comparison to the literature

We used the sagittal plain radiographs of 60 healthy subjects, available from previous retrospective studies, 42% men and 58% women, mean age 43 years (range 22–68 years). Validity was estimated for sagittal postural X-ray film analysis by comparing our measurements to published reference values for main pelvic and spine parameters.

¹once agreement between observers and normal distributions have been checked

Limits

The last step of our study was the evaluation of the software's limits for clinical quantitative analyses in extremely difficult situations, such as severe deformities. Two observers analysed 60 plain radiographs (30 sagittal and 30 frontal) of patients suffering from severe scoliosis (Cobb angle $> 40^\circ$) with different film qualities.

A Wilcoxon test was performed to determine the existence of a statistically significant difference between the values obtained by the two observers. Then we calculated the inter-observer reproducibility coefficient ($1.96 \times \text{SD}$ of the inter-observer errors).

Results

Precision tests (dynamic modules)

There were no statistically significant differences between the two methods: the reference and Spineview calculation of ROMs. (P -value: 0.041).

No significant differences were found either for the ROMs calculated at different vertebral levels. Moreover, a normal distribution was found for the global error. Therefore, we calculated a global precision estimator both for cervical and lumbar dynamic module. The results are presented in Table 1.

Reproducibility of measurements

Taking into account the different parameters and types of X-ray films analyzed, the results of the reproducibility study will be presented separately for dynamic and postural modules.

Dynamic modules

Lumbar dynamic module: no statistically significant differences were found between the ROM calculated at

Table 1 Results of precision tests for dynamic modules

Module	Number of analyzed items	Mean error (°)	Standard deviation ^a (SD)	Precision estimator (°)
Cervical	64	-0.06	1.4°	± 2.8
Lumbar	25	-0.03	1.0°	± 2.0

^a Standard deviation of the errors (Spineview 2.0 versus reference)

different vertebral levels, except for L5S1 and L1L2, as shown in the Table 2. Except for these two levels, there are no statistically significant differences between ROM variances at each level; this allowed us to calculate a global CR, which is 1.6°. We mention that CR considered for the L5S1 and L1L2 levels is 2.6°.

Mean reproducibility error for the linear parameters is ±0.8 mm for the vertebral translations and ±1.6 mm for the vertebral height, except for the L5S1 level, where reproducibility of the vertebral height is ±4 mm; for discal height the reproducibility estimator is 1.8 mm.

Cervical dynamic module: we found no statistically significant differences between the ROM calculated at different vertebral levels (variance ranging from C3C4 to C6C7 between 0.6° and 1.1°), except for C7T1 level (1.7°). Therefore, the global reproducibility error (for a 95% confidence interval) was estimated to ±2.2° for the C6–C7 to C3–C4 levels and to ±3.4° for the C7T1 level.

Mean reproducibility errors for the length parameters are ±1.4 mm for the vertebral translations and ±1 mm for the vertebral and discal height.

Postural modules

The ANOVA analysis of the results showed no statistically significant difference between the values measured by the three observers; still, for lumbar Cobb's angle, lumbar lordosis, sacral slope and some vertebral parameters calculated at S1, T3, T4 and some cervical levels (i.e. C5C6 and C2C3) we are at the limits of agreement (both for Tukey HSD and Fisher LSD tests with a confidence range of 95%).

As a complementary test, the intra-class correlation coefficient CCI [21, 28, 37] for the calculated parameters were evaluated. A good correlation was found in within-subject repetitive measures (CCI = 0.99–1.00) and good agreement between observers (CCI = 0.986–0.999).

The analysis of variance for each parameter corresponding to a vertebral level (e.g., intervertebral angle T3T4 or vertebral wedging of L2) showed some particularities, which will be presented for each postural module. We also present in the following the global reproducibility coefficients (all observers, all film qualities) of the most important parameters calculated by the sagittal and frontal modules.

Sagittal module The global CR ranges calculated for *spinal and pelvic parameters* are for the angles: 0.1/5.5°, for the linear parameters: 1/4 mm and for those representing ratios: 3.2/8.1. The 95% CI is described by the measured value ±CR. The reproducibility estimators obtained for a routine quality of X-ray film are presented in Table 3.

Vertebral parameters The analysis of variances showed a clear difference between *angular* parameters (i.e. intervertebral angles, vertebral wedging) calculated at most levels, with CR of 3° and those calculated at L5S1, T6T7, T1T2 and cervical levels, where the CR may reach 4.5°. The same results were found for indexes, with global CR of 5 and 7%, respectively. For the *linear* parameters (vertebral and discal height) the global CR is 2 mm (except for the parameters calculated for T3-T5 levels, where CR may reach 4 mm).

Frontal module The global reproducibility estimators for *main spinal and pelvic parameters* are about ±3° for the pelvic obliquity, ±4° for the Cobb angles and about ±1.3° for the global inclination. Furthermore, for a routine quality of X-ray film, the reproducibility estimators (inter-observer) calculated for same parameters were, 1.8°, 3° and 0.6° respectively.

Vertebral parameters Similarly to the sagittal X-ray films analysis, for most levels, global reproducibility error was equal to ±3° for the angles (i.e. intervertebral angles and vertebral wedging), to ±2 mm for linear

Table 2 Analysis of variance for lumbar dynamic module—reproducibility tests

Sample	Sample size	Variance	SD	Reproducibility estimator ^a
SpineAngL5Sacrum	90	1,7	1.3	2.6
SpineAngL4L5	90	0,7	0.9	1.6
SpineAngL3L4	90	0,5	0.7	
SpineAngL2L3	90	0,7	0.7	
SpineAngL1L2	90	1.6	1.3	2.6

^a Reproducibility coefficient at 95% CI

Table 3 Descriptive statistics for reproducibility of parameters calculated on a routine quality sagittal X-ray film

Parameter	C1C7 lordosis (°)	T4T12 kyphosis (°)	L1S1 lordosis (°)	TLPL PRT12 ^a (°)	T1 sagittal offset (°)	T9 sagittal offset (°)	Global inclination (°)	EAM ^b overhang (mm)	T9 overhang (mm)	Sacral slope (°)	Pelvic version (°)	Incidence (°)	Pelvic angle (°)	PRSI angle (°)
Reproducibility estimator	4.4	5.2	4.4	2.4	0.1	0.2	0.1	1.5	1.6	4.0	1.2	3.4	1.4	3.0
CV ^c (SD/mean)	0.06	-0.03	0.03	0.01	-0.01	-0.01	-0.01	-0.01	-0.09	0.05	1.25	0.04	0.13	0.04

^a Total lumbopelvic lordosis PR-T12 ^b External auditory meati ^c Coefficient of variation

parameters (i.e. vertebral and disc's height) and $\pm 6\%$ for the listhesis (ratio). For T2-T4 and T9-T11 levels, CR calculated was 5° for the angles and 11% for the ratios. Also, for linear parameters calculated at T2-T4, T8-T12 and L5S1 levels the reproducibility estimator equals 3 mm.

Validity

The values obtained for the main parameters in our healthy population compared to the literature are presented in the Table 4.

Limits

For the extreme cases of severe adult scoliosis, the statistic tests (Wilcoxon, sign, Student) showed no significant difference between the two observers.

For both sagittal and frontal modules, the reproducibility ranges are presented in the Table 5. The analysis of the reproducibility estimators showed that errors in the calculation of *spinal parameters* increase of 100% for kyphosis, lordosis and Cobb angles, which are less precise (8°) than those calculated for patients suffering from degenerative disease ($3-4^\circ$) and also for T1 and T9 sagittal offsets, which still have a good reproducibility ($0.6-0.8^\circ$).

Reproducibility of *pelvic parameters* measurements diminishes of 50% for sacral slope, pelvic version and incidence and of 100% for pelvic obliquity. However, the reproducibility estimators of some parameters, like pelvic version and pelvic angle are about 3° .

For the *vertebral parameters* (intervertebral angles, discal heights and listhesis) reproducibility diminished of 50% for the frontal module and only of 20% for the sagittal analysis, with regard to the values we calculated for degenerative patients.

Discussion

The main purpose of our study was to evaluate the Spineview 2.1 software, dedicated to clinical quantitative analysis. In order to test its reliability, validity and limits in clinical routine conditions, several qualities of patients X-ray films were chosen.

Knowing that small but important changes in a patient's evolution may require investigation in clinical routine activities, as well as in research studies, the first condition to perform a reliable quantitative analysis is the reliability of measurements, defined in the literature [7, 14] as the measure of the extent to which a test is reproducible over different situations. Its evaluation involves precision and reproducibility tests.

Table 4 Comparison to the literature for reference values of main spinal and pelvic parameters obtained from asymptomatic subjects

Parameter	Number of subjects	Thoracic kyphosis (T1-T12)	Kyphosis T4/T12	Lordosis L1/S1	Lumbar lordosis (L1-L5)	T9 sagittal offset	Global inclination	Sacral slope	Pelvic version	Pelvic incidence	TLPL PRT12 ^c	PRS1 angle	Pelvic angle
Duval-Beaupère[4]	17		34 ± 16					41 ± 14	11 ± 12	52 ± 18			
Guigui [6]	250		31 ± 18	61 ± 26		11 ± 6		44 ± 16	13 ± 12	55 ± 22	93 ± 22	32 ± 20	18 ± 12
Jackson [8, 9]	75	46 ± 22	39 ± 24	62 ± 22									
Legaye [11]	38	^b 46 ± 20		^b 61 ± 22		11 ± 6							
Marty [12]	44			^b 60 ± 20				42 ± 16	12 ± 14	53 ± 22			
Rainics [17]	30	39 ± 22			55 ± 12	11 ± 6	1 ± 6	41 ± 18	11 ± 12	51 ± 22			
Rillardon [18]	100	40 ± 18			43 ± 24	10 ± 6		43 ± 12	11 ± 12	54 ± 18			
Stagnara [22]	100							42 ± 16	13 ± 14	55 ± 16			
Templier [23]	13			50 (32 ÷ 84)				41 (19 ÷ 65)					
Vaz [26]	100	^b 47 ± 18			38 ± 26			40 ± 16	15 ± 14	55 ± 24			
Voutsinas [27]	670 ^a			57 ± 18				39 ± 18	12 ± 12	52 ± 24			
Our study	60	46 ± 18	39 ± 16	57 ± 22	49 ± 18	11 ± 6	0 ± 6	39 ± 16	13 ± 14	51 ± 22	92 ± 18	33 ± 22	18 ± 14

^a Adolescents^b Most tilted vertebrae^c Total lumbopelvic lordosis PR T12[8]

Values are mean ± 2SD

Table 5 Reproducibility ranges for main postural parameters analysing severe scoliosis radiographs

Parameter	CR ranges	Examples
Angles (°)	0.6 – 8.0	T1 sagittal offset: 0.6° Pelvic angle: 2.9° T4T12 kyphosis: 8.0°
Distances (mm)	3.3 – 8	Disc's height: 3.3 mm T9 overhang: 8 mm
Other (%)	9–16	Sagittal listhesis: 9% Frontal listhesis: 16%

Precision

We have to underline the fact that the announced values represent an estimator for the measurement precision using this software versus “in vitro” referential (metallic rods). A previous study showed that axial rotation inferior to 20° does not yield significant changes in the angular measurements (for a 0° incidence of the X-ray beam) [23]. Focus height could introduce slight bias; therefore, we followed the generally used focus, i.e. L3 vertebra for lumbar spines and C4 vertebra for the cervical one, as mentioned in Materials and methods. The precision of angular parameters measurement for dynamic modules was established at ±2° for the lumbar region analysis and ±3° for the cervical region. This difference might be due to the different geometry between the cervical and lumbar vertebrae.

These values may also be compared to those established by the precision testing of Spineview 1.0 software for the lumbar dynamic module [17, 23] which were quite similar (±1.5°). The original X-ray films of the mentioned study were scanned again for the present one and film degradation could be an explanation for the slight difference observed.

Reproducibility

The comparison of the intra- and inter-observer measurements agreement by means of various statistical tests allowed us to further apply descriptive statistics, in order to quantify the reproducibility errors. Among common descriptors like mean, standard deviation, SD, coefficient of variation, Cv, we calculated the coefficient of reproducibility, CR, which is defined [3] as the expectation that 95% successive measures of a known parameter using Spineview 2.1 software would differ one from another by less than two standard deviations (CR = 2 SD).

Dynamic modules

We found a significant difference in the variances calculated for the ROM corresponding to the extreme

levels, with regard to the other levels. For example, variance of L5S1 and L1L2 is different from that calculated for the other lumbar levels; dynamic radiographs being focused on L3 region, the higher reliability of measurements corresponds to this region and diminishes for the other levels according to the distance.

Calculation of ROM at most levels is highly reliable, the reproducibility estimator being comparable to the precision one. For the extreme levels (L5S1, L1L2 and C7T1*) reproducibility is lower than precision; the observed difference (0.4 and 0.6°) might be related to the general better visibility of the spine segments compared to patients' vertebrae on X-ray films.

Therefore, the investigations on the ROM calculation showed that the reproducibility errors are in agreement with the announced precision estimators.

Concerning the other parameters calculated by dynamic modules reproducibility of distances linear parameters seems to depend on the geometry and motion of the vertebra. This might explain why vertebral sizes are measured with a reproducibility estimator of 1 mm for the cervical levels and of 1.6 mm for the lumbar ones; meanwhile the translation reproducibility is respectively of 1.4 and 0.8 mm.

Postural modules

The results of the present study showed a mean reproducibility error of 4° for the angular parameters and of 3 mm for the distance parameters. We found a strong inter-observer agreement and a good correlation for the within-subject calculated values.

Spinal and pelvic parameters calculation depends on the visibility of anatomic landmarks. Cervical and lumbar lordosis are calculated with a CR of 4–4.4°; the total lumbopelvic lordosis TLPL PRT12 seems to be more reproducible (CR = 2.4°). The kyphosis CR varies between 2.2° and 5.2° according to the visibility of the T1–T4 region.

Spinal parameters for which calculation is based on several anatomic landmarks, such as T1, T9 sagittal offsets, global inclination, T9 and EAM overhang, showed good reproducibility (0.1°–0.6° and respectively, 1.5 mm).

Pelvic parameters such as sacral slope, incidence and PRS1 angle have a CR of about 4°; meanwhile, pelvic version and pelvic angle show good reproducibility: CR = 1.2–1.4°. Unlike Rillardon et al. [18], we found no major difference for the reproducibility of pelvic parameters defined by Duval-Beaupère et al. [4] and Legaye et al. [11] versus those defined by Jackson et al. [8, 9].

The vertebral parameters variance, depending on the visibility of anatomic landmarks, increases for the mid thoracic, cervical and L5S1 levels (sagittal analysis) and for T2–T4 and T9–T11 levels (sometimes also L5S1) for the frontal module.

Therefore, the quality of the image is of paramount importance. Knowing that SpineView 2.1 software allows for identification and calculation of parameters on the whole spine as well as on selected spinal segments, this enables the investigator to choose his area of analysis, for example a zone of good visibility, and to take into account only the parameters relative to this region. As the quality of the X-ray films is supposed to increase with the new digital systems, our study for increasing the robustness and the precision of the software continues.

When looking to the literature one can note that few studies are comparable to ours. Manual measurement reproducibility for some angular parameters is sometimes documented [2, 5, 8, 9, 27] and the results are comparable to ours. As an example, we would like to mention the reproducibility of TLPL PR T12 (total lumbopelvic lordosis PR T12), which was assessed by Jackson to 3° and by our study to 2.4°; in comparison to a CR of 4° for classical lumbar lordosis, the calculation of this parameter seems to be more reliable.

Few studies assess the reproducibility of X-ray films based spinal and pelvic parameters measurements using computer-assisted methods; we present in Table 6 a comparison between reproducibility estimators obtained for Cobb angle, T4T12 kyphosis, lumbar (L1L5) and L1S1 lordosis, T1 and T9 sagittal offsets, global inclination, and pelvic parameters (literature versus present study).

Shea et al. [20] and Cheung et al. [3] obtained similar results for the intra-observer reproducibility of Cobb angles; we evaluated the reproducibility for the same parameter at 3° (intra-observer) and 4° (inter-observer), which is consistent with their findings. Furthermore, Cheung et al. announced a reproducibility of 3.3° for the T4T12 kyphosis (intra-observer evaluation), which is highly depending on the X-ray film quality; we found the reproducibility estimator of 2.2° for this parameter for a good quality of X-ray films but the inter-observer reproducibility for routine radiographs quality is about 5°.

Spinal parameters such as kyphosis and lordosis are calculated with a reproducibility of 4–5°; it is stated in the medical literature that the variability in assessing high thoracic and low lumbar anatomical landmarks is strongly related to their visibility on the X-ray films. Our results are comparable to those announced by Rajnics et al. [17] and Rillardon et al. [18], the difference coming from the different type of X-ray films analyzed. The mentioned studies calculated spinal and pelvic parameters on asymptomatic subjects' radiographs. Furthermore, Rajnics et al. used a protocol of successive measurements, which facilitates the comparison; main differences in the reproducibility of parameters were found for T4T12 kyphosis (1.3°) and global inclination (1°).

Rillardon et al. measured 100 X-ray films of asymptomatic subjects in intra and inter-observer analysis; the

Table 6 Reliability of the version 1.0 [17] and 2.0 of Spineview in terms of reproducibility

Reproducibility estimator ^a	Parameter (°)							
	Cobb angle°	T4T12 kyphosis	Lumbar lordosis	T9 sagittal offset	Global inclination	Sacral slope	Pelvic version	Incidence
Shea et al. [20] intra observer	2.6							
Cheung et al. [3] intra observer	3.1	3.3						
inter observer	4.8	5.4						
Rajnic et al. [17] inter observer		6.5	4.3	0.3	1.1	4.1	1.5	3.8
Rillardona et al. [18] inter observer		4.2	3	2		2.4	2	2.4
Our study inter observer	4	5.2	4.6	0.2	0.1	4.0	1.2	3.6

^aThese reproducibility estimators were extrapolated from their measurements

reproducibility estimators calculated are a little better than those we announced for kyphosis, lordosis, sacral slope and incidence (the maximum difference is 1.6°) and a little less accurate for T9 sagittal offset (same difference). We suppose that the observed difference might come from the higher visibility of vertebral endplates on the X-ray films for asymptomatic subjects versus patients and also from the difference of the statistical methods and test size.

As a conclusion, we found our measurements comparable to those published in the literature; nevertheless, our research for increasing precision and robustness of the analysis continues.

The second part of the study focused on a practical application in clinical environment. Reliability by itself is not a sufficient estimator because reliable results might be biased (containing systematic errors).

Validity [7, 14] refers to the question whether the measure is assessing what is intended. We found that our reference values on the asymptomatic subjects are comparable (Table 4) to those published by several authors. For the parameters well defined in the literature such as T4T12 kyphosis, T9 sagittal offset and pelvic parameters (i.e. sacral slope, pelvic version, incidence, pelvic and PRS1 angle) we found a very good agreement between our values and the literature, except for the L1S1 lordosis where our reference value is comparable to that announced by Voutsinas et al. [27] and inferior (3°) to that announced by other authors. This difference might be related to the different reliabilities and methods of measurement, different mean ages of the populations and also to the variability of sacrum identification.

Thoracic kyphosis and lumbar lordosis are less clearly-defined; most authors calculate these angles between the upper and respectively lower plates of the most tilted vertebrae of the sagittal shape of the spine. For the thoracic region, these angles are calculated most

often between T1 and T12, with a good agreement between different studies (Table 4). Lumbar lordosis, which is calculated theoretically between the most tilted vertebrae, is often considered either between the lower plate of T12 and that of L5 or the sacral plate, either between L1 and L5. In our population of 60 asymptomatic subjects, we found that one-fifth of the lumbar lordosis values were calculated between L2 and L5 and the other four-fifth between L1 and L5, which confirms the hypothesis of Stagnara et al. (the value we compared to the literature is the value obtained for L1L5 lumbar lordosis).

Therefore, we might consider that our reference values match those published in the literature for main spinal and pelvic parameters.

Limits

When analysing severe scoliosis on various qualities of X-ray films, manual measurements can be difficult and the results very biased. Computer-assisted measurements might be helpful but reliability of analyses in such extreme conditions must be evaluated.

Our study showed that the visibility of anatomic landmarks seems to depend not only on the quality of X-ray films, but also on the extent of the deformation; as an example, the areas outlined in Fig. 2 show how problematic a landmark identification can be in some cases. Therefore, the reproducibility of parameters was highly variable:

- In areas of good visibility, the CRs calculated for frontal and sagittal modules are comparable to those obtained in routine studies (examples: T1 and T9 sagittal offset, pelvic angle and listhesis still have a good reproducibility),
- Whereas in cases of low visibility (sacrum, lumbar and mid thoracic levels), the reproducibility error can

Fig. 2 Examples of X-ray films analyzed

double (example: T4T12 kyphosis, reproducibility error passes from 4° to 8°). Some concerned parameters become less reliable, i.e. kyphosis, Cobb angle and lordosis; others, such as pelvic angle ($CR = 1.4^\circ$ (patients suffering from degenerative diseases) to 2.9° (severe scoliosis)) still present good reproducibility estimator.

The other drawback is two-dimensional analysis in situations in which the three-dimensional deformity yields major bias. Therefore, precaution must be taken using Spineview 2.0 software in such analyses.

Conclusions

This study allowed to evaluate SpineView 2.1 software and to estimate the reliability and validity of clinical measurements. The precision estimators calculated for inter-vertebral mobility are $\pm 2^\circ$ for the lumbar analysis and $\pm 3^\circ$ for the cervical one. Reproducibility ranges between 0.1° and 5.5° for angular parameters and between 1 mm and 4 mm for linear parameters.

The reliability of measurements highly depends on the visibility of anatomic landmarks on the X-ray film. Limits of the software were also highlighted, in some pathologic conditions such as severe adult scoliosis where visibility is usually poor and three-dimensional deformity may yield major bias.

Fast and reliable computer-aided spinal X-ray films quantitative analyses may significantly contribute to monitor small but clinically important changes in the evolution of a patient and constitute the first steps towards outcomes assessment, essential for developing evidence based medicine. Such tools open the way to extensive standardized measurements in different medical care centres (as far as the radiographic protocol is identical) providing the investigators with a large amount of collected data by multicentric evaluation and comparison between studies.

Acknowledgements The authors acknowledge H. Carrier, A. Raoult, who participated to inter-observer measurements, A. Jacotot, V. Pomeroy and I. Kojadinovic who helped us managing the statistical tests and A. Mitulescu, who kindly offered her support in editing and revising this paper. Note: Spineview software was developed by the teams of LBM& LIO, in collaboration with the Surgiview company (Paris, France).

References

1. Altman D, Bland M (1983) Measurement in medicine: the analysis of method comparison studies. *Statistician* 32:307–317
2. Bernhardt M, Bridwell KH (1989) Segmental analysis of the sagittal plane alignment of the normal thoracic and lumbar spines and thoracolumbar junction. *Spine* 14(7):717–721
3. Cheung J, Wever DJ, Veldhuizen AG, Klein JP, Verdonck B, Nijluning R, Cool JC, Van Horn JR (2002) The reliability of quantitative analysis on digital images of the scoliotic spine. *Eur Spine J* 11:535–542
4. Duval-Beaupère G, Schimt C, Cosson PH (1992) A barycentrometric study of the sagittal shape of spine and pelvis. *Ann Biomed Eng* 20(4):51–62
5. Gelb DE, Lenke LG, Bridwell KH, Blanke KB, McEnery KW (1995) An analysis of sagittal spinal alignment in 100 asymptomatic middle and older aged volunteers. *Spine* 20(12):1351–1358
6. Guigui P, Levassor N et al (2003) Physiological value of pelvic and spinal parameters of sagittal balance: analysis of 250 healthy volunteers. *Rev Chir Orthop Reparatrice Appar Mot* 89(6):496–506
7. Hopkins WG (2000) Measures of reliability in sports medicine and science. *Sports Med* 30:1–15
8. Jackson RP, Kanemura T, Kawakami N, Hales C (2000) Lumbopelvic lordosis and pelvic balance on repeated standing lateral radiographs of adults volunteers and untreated patients with constant low back pain. *Spine* 25(5):575–586
9. Jackson RP, McManus AC (1994) Radiographic analysis of sagittal plane alignment and balance in standing volunteers and patients with low back pain matched for age, sex and size. *Spine* 19(14):1611–1618
10. Kauffmann C, Godbout B, De Guise J (1998) Simplified active contour model applied to bone structures segmentation in digital radiographs. In: *Medical imaging 1998, SPIE'S International Symposium*, San Diego, California, USA, 21–27 February, 3338: 663–672
11. Legaye J, Hecquet J, Marty C, Duval-Beaupère G (1993) Equilibre sagittal du rachis. *Rachis* 5(5):215–226
12. Marty C, Boisaubert B, Descamps H, Montigny JP, Hecquet J, Legaye J, Duval-Beaupère G (2002) The sagittal anatomy of the sacrum among young adults, infants, and spondylolisthesis patients. *Eur Spine J* 11(2):119–125
13. Percy MJ, Bogduk N (1988) Instantaneous axes of rotation of the lumbar intervertebral joints. *Spine* 13(9):1033–1041
14. Portney LG, Watkins MP (1993) *Foundations of clinical research Applications and practice*. Appleton & Lange, Norwalk ISBN 0-8385-1065-5, pp 509–516
15. Precision measurement of disc height, vertebral height and sagittal plane displacement from lateral radiographic views of lumbar spine (1997) *Clinical biomechanics* 12(suppl):1
16. Prothèses discales et arthrodèses dans la pathologie dégénérative du rachis lombaire - rapport de l'ANAES, Mai 2000
17. Rajnics P, Pomero V, De Guise J, Mouilleseaux B, Lavaste F, Skalli W (2001) Computer-assisted assessment of spinal sagittal plane radiographs. *J Spinal Disord* 14:135–142
18. Rillardon L, Levassor N, Guigui P, Wodecki P, Cardinne L, Templier A, Skalli W (2003) Validation of a tool to measure pelvic and spinal parameters of sagittal balance. *Rev Chir Orthop Reparatrice Appar Mot* 89(3):218–227
19. Sai Wing Lee, Kris WN Wong et al (2002) Development and validation of a new technique for assessing lumbar spine motion. *Spine* 27:E215–E220
20. Shea KG, Stevens PM, Nelson M, Smith Jt, Masters KS, Yandow S (1998) A comparison of manual versus computer-assisted radiographic measurement. Intraobserver measurement variability for Cobb angles. *Spine* 23:551–555
21. Shrout P, Fleiss J (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 68:420–428
22. Stagnara P, De Mauroy JC, Dran G et al (1982) Reciprocal angulation of vertebral bodies in a sagittal plane: approach to references for evaluation of kyphosis and lordosis. *Spine* 7(4):335–342
23. Templier A, Skalli W, Diop A, Vital JM, Mazel C, Saillant G, Delecrin J, Dubousset J, Lavaste F (1999) Evaluation radiographique quantitative des arthrodèses rachidiennes lombaires: protocole et analyse préliminaire d'une série multicentrique de 119 cas. *Rachis*, pp 23–37
24. Toussignant M, Duclos E et al (2002) Validity study for the cervical range of motion device used for lateral flexion in patients with neck pain. *Spine* 27:812–817
25. Troyanovich SJ, Cailliet R, Janik TJ, Harrison DD, Harrison DE (1997) Radiographic mensuration characteristics of the sagittal lumbar spine from a normal population with a method to synthesize prior studies of lordosis. *J Spinal Disord* 10:380–386
26. Vaz G, Roussouly P, Berthonnaud E, Dimnet J (2002) Sagittal morphology and equilibrium of pelvis and spine. *Eur Spine J* 11(1):80–87
27. Voutsinas SA, MacEwen GD (1986) Sagittal profiles of the spine. *Clin Orthop* 210:235–242
28. White Paper on Sagittal Plane Alignment, Scoliosis Research Society