The Spine Journal 14 (2014) 274–281

2013 Outstanding Paper Runner-up

Reliability of computer-assisted lumbar intervertebral measurements using a novel vertebral motion analysis system

Matthew S. Yeager, BS[a], Daniel J. Cook, MS[a], Boyle C. Cheng, PhD[a,b,*]

[a]*Department of Neurosurgery, Allegheny General Hospital, 420 East North Ave., Suite 302, Pittsburgh, PA 15212, USA*
[b]*Department of Neurosurgery, Drexel University College of Medicine, 420 East North Ave., Suite 302, Pittsburgh, PA 15212, USA*

Received 21 January 2013; accepted 19 October 2013

Abstract

BACKGROUND CONTEXT: Traditional methods for the evaluation of in vivo spine kinematics introduce significant measurement variability. Digital videofluoroscopic techniques coupled with computer-assisted measurements have been shown to reduce such error, as well as provide detailed information about spinal motion otherwise unobtainable by standard roentgenograms. Studies have evaluated the precision of computer-assisted fluoroscopic measurements; however, a formal clinical evaluation and comparison with manual methods is unavailable. Further, it is essential to establish reliability of novel measurements systems compared with standard techniques.

PURPOSE: To determine the repeatability and reproducibility of sagittal lumbar intervertebral measurements using a new system for the evaluation of lumbar spine motion.

STUDY DESIGN: Reliability evaluation of digitized manual versus computer-assisted measurements of the lumbar spine using motion sequences from a videofluoroscopic technique.

PATIENT SAMPLE: A total of 205 intervertebral levels from 61 patients were retrospectively evaluated in this study.

OUTCOME MEASURES: Coefficient of repeatability (CR), limits of agreement (LOA), intraclass correlation coefficient (ICC; type 3,1), and standard error of measurement.

METHODS: Intervertebral rotations and translations (IVR and IVT) were each measured twice by three physicians using the KineGraph vertebral motion analysis (VMA) system and twice by three different physicians using a digitized manual technique. Each observer evaluated all images independently. Intra- and interobserver statistics were compiled based on the methods of Bland-Altman (CR, LOA) and Shrout-Fleiss (ICC, standard error of measurement).

RESULTS: The VMA measurements demonstrated substantially more precision compared with the manual technique. Intraobserver measurements were the most reliable, with a CR of 1.53 (manual, 8.28) for IVR, and 2.20 (manual, 11.75) for IVT. The least reliable measurements were interobserver IVR and IVT, with a CR of 2.15 (manual, 9.88) and 3.90 (manual, 12.43), respectively. The ICCs and standard error results followed the same pattern.

CONCLUSIONS: The VMA system markedly reduced variability of lumbar intervertebral measurements compared with a digitized manual analysis. Further, computer-assisted fluoroscopic imaging techniques demonstrate precision within the range of computer-assisted X-ray analysis techniques.

Keywords: Reliability; Lumbar; Computer assisted; Intervertebral measurement; Videofluoroscopy; Vertebral motion analysis

---

FDA device/drug status: Approved (KineGraph Vertebral Motion Analyzer).

Author disclosures: ***MSY:*** Nothing to disclose. ***DJC:*** Trips/Travel: Innovative Performance Technologies (A). ***BCC:*** Consulting: Total Spine Solutions (C); Scientific Advisory Panel Board of Directors: Orthokinematics (C, shares); Grants: Medtronic (F, Paid directly to institution), Applied Spine (E, Paid directly to institution), Rachiotek (E, Paid directly to institution), Alphatec (F, Paid directly to institution), Scient'x (F, Paid directly to institution), IMDS (F, Paid directly to institution), Olympus (F, Paid directly to institution), Globus (D, Paid directly to institution), Stryker (F, Paid directly to institution).

The disclosure key can be found on the Table of Contents and at www.TheSpineJournalOnline.com.

* Corresponding author. Department of Neurosurgery, Allegheny General Hospital, 420 East North Ave., Suite 302, Pittsburgh, PA 15212, USA. Tel.: (412) 359-4020; fax: (412) 359-4811.

E-mail address: boylecheng@gmail.com (B.C. Cheng)

1529-9430/$ - see front matter
http://dx.doi.org/10.1016/j.spinee.2013.10.048

Introduction

Imaging modalities such as standard roentgenograms (X-rays), computed tomography, and magnetic resonance imaging have become essential to the evaluation of bone and soft tissue in patients with spinal pathologies. These techniques capture valuable static images of the spine, yet lack the capability of providing detailed information about spinal motion. Dynamic end-range X-rays, the standard for assessing range of motion (ROM) and vertebral translation are typically evaluated by measuring intervertebral rotation and translation (IVR and IVT) through manual identification of vertebral margins using a ruler and protractor or digital techniques involving medical imaging software. However, the error associated with these methods is high and within the range of currently proposed motion guidelines [1–5].

Current thresholds for indication of successful fusion, motion preservation, and instability are not well defined. The US Food and Drug Administration defines successful fusion for Investigational Device Exemption trials as less than 5° IVR and less than 3 mm IVT in the sagittal plane, although others have reported upper IVR thresholds anywhere from 1° to 4° [6–8]. Successful motion preservation for total disc replacement has been reported as low as 2° to 3°, obviously well within the range of successful fusion guidelines [9,10]. Reimbursement guidelines for instability (InterQual, McKesson, San Francisco, CA, USA) suggest IVR greater than 22° and IVT greater than 3 mm as unstable, whereas the American Medical Association indicates bounds anywhere from 15° to 25° and 4.5 mm [11]. Lower thresholds of 10° IVR and 4 mm IVT are also commonly accepted to infer instability [12,13]. Given the wide range of proposed motion guidelines and arbitrary definitions for quantification of instability, a better understanding of spinal motion characteristics through standardized, accurate, and reliable functional analysis is clearly necessary.

Techniques utilizing landmark verification protocols, computerized image processing software, and automatic vertebral tracking algorithms have been evaluated that demonstrate more accurate and reliable intervertebral motion measurements in functional radiographic images [14–19]. These methods utilize either plane film radiographs or digitized videofluoroscopy (DVF) to capture images of the lumbar spine, the latter of which is capable of capturing spinal motion with less radiation exposure than that of traditional end range X-rays. Prior studies have evaluated accuracy and reliability of computer-assisted DVF measurements and have demonstrated errors comparable to those obtained from computer-assisted X-ray techniques. These studies have reported IVR accuracy and reliability errors in the range of 0.13° to 1.18° (standard deviation) [15–20].

To the authors' knowledge, only one study to date has formally evaluated and compared intra- and interobserver reliability of automated versus manual measurements. Pearson et al. [5] compared the Quantitative Motion Analysis system, today's clinical standard in dynamic plane film analysis, with a digitized manual technique. The results support conclusions that computer-assisted processing methods significantly improve intervertebral motion measurements. However, no such evaluation is available for DVF techniques. Further, no automated system to date has been formally evaluated that attempts to control initial sources of variability in clinical radiographs by means of standardizing patient bending angles and image acquisition techniques. A system described by Breen et al. [21] used DVF with motorized recumbent patient bending platforms to demonstrate maximum errors (root mean square) of under 2°; however, a formal reliability analysis was not completed.

A new system for evaluation of spinal motion has recently been approved by the FDA for commercial use, utilizing upright and recumbent patient bending platforms with DVF and automated vertebral tracking algorithms (KineGraph Vertebral Motion Analysis [VMA], OrthoKinematics, Austin, TX, USA). This study assesses intra- and interobserver reliability of this system using a prospective analysis of retrospectively collected image data.

Materials and methods

Image acquisition

The VMA system utilized a combination of upright and recumbent controlled patient bending platforms, a standard 12-inch surgical C-Arm (OEC 9800 Series, General Electric, Fairfield, CT, USA), and an adjacent computer-mounted console equipped with data acquisition hardware (Accustream Express As205A, Foresight Imaging, Chelmsford, MA, USA) and proprietary control software. The independent bending platforms consisted of a radiolucent disc, which acted as the center of rotation, while adjustable components accommodated for varying physical patient characteristics and bolsters secured the pelvis to isolate trunk bending as the torso completed a predetermined total ROM of 70° at a rate of approximately 5° per second.

The upright motion platform guided active lumbar bending, under the weighted condition, while constricting flexion extension (FE) to the sagittal plane. The recumbent platform controlled passive lumbar bending, which minimized the gravitational and muscular forces that present during standing radiographs. Patient positioning on each platform can be visualized in Fig. 1. FE angles spanned a range of ±35° for recumbent motion, and the ROM presets were adjusted to allow for 20° of extension and 50° flexion for upright motion. This compensated for the reduced capability of lumbar extension resulting from extended hips in an erect posture.

Upon initiation of the test movement, a fluoroscopic sequence of lumbar motion was captured at 8 pulses per second. Flexion and extension were captured as separate sequences, which began in a neutral position, progressed to the predefined maximum angle, and returned to neutral. The console PC grabbed real-time images from the C-Arm analog video output port, which were digitized at a depth of

Fig. 1. Upright (Left) and recumbent (Right) patient bending platforms with patient positioned for flexion–extension imaging.

8 bits, sampled at 8 frames per second, and stored in Digital Imaging and Communications in Medicine format for future processing.

Experiment design and image analysis

This study combines retrospectively collected image data from 2 separate multicenter studies that utilized the VMA system. All centers (n=7) employed the same imaging protocol described. An aggregated image set was created from 61 subjects (48% female, 52% male; mean age, 46.5±15.5 years) classified as either asymptomatic (n=34), preoperative with a confirmed pathology (n=14), or postoperative with a previous lumbar procedure (n=13). For each patient, a single lateral image sequence of flexion or extension was selected from either the upright or recumbent position. The aggregated image set, composed of 205 intervertebral levels, was then processed to generate IVR and IVT measurements using two different techniques: automated measurements with VMA tracking algorithms and a manual technique with medical imaging software (ImageJ, National Institutes of Health, Bethesda, MD, USA).

The automated tracking process consisted of a template-based approach using normalized, mean-centered cross-correlation as a matching criterion to automatically identify each vertebra throughout the image sequence. This method has been previously described [16,21]. Although measurements were generated for every frame in the sequence, only measurements at neutral and maximum patient bending angles were used. The same two image frames used for these measurements were digitally reproduced for the manual process, which consisted of a standard ‘‘point and click’’ method to define vertebral margins and generate measurements in ImageJ. The IVR for all measurements was defined as the change in angle between the inferior endplate of the superior vertebra and the superior endplate of the inferior vertebra in the motion segment. The IVT was defined using the Stokes and Frymoyer [22] method and expressed as a percentage of vertebral body depth.

A set of three physicians individually processed the VMA measurements, and a separate set of three physicians individually conducted the manual analysis. Each observer completed their assigned measurement technique twice, the second of which was separated by at least 2 weeks duration. Mean angles for all IVR and IVT measurements were calculated, as well as the amount that each level (L1–L2 to L5–S1) contributed to the data. Intra- and interobserver datasets where then compiled to produce descriptive reliability statistics.

Statistical analysis

The Bland-Altman method was employed to evaluate the agreement within and between observers. The differences between each observer’s first and second round

Table 1
Average angles and translations measured for both computer-assisted and manual techniques

	IVR (°)		IVT (%)	
Technique	Flexion	Extension	Flexion	Extension
VMA	4.40	2.00	2.10	1.13
DM	3.95	1.53	0.23	0.84
Difference	0.45	0.47	1.87	0.29

IVR, intervertebral rotation; IVT, intervertebral translation; VMA, computer-assisted vertebral motion analysis; DM, manual vertebral motion analysis.

Table 2
Contributions of intervertebral levels comprising the dataset

Level	n	Contribution (%)
L1–L2	14	7
L2–L3	37	18
L3–L4	58	28
L4–L5	51	25
L5–S1	45	22
Total	205	

Table 3
Bland-Altman and Shrout-Fleiss statistics

Bland-Altman								ICC		
Measurement	Mean of differences	CI	Upper LOA	CI	Lower LOA	CI	CR	ICC (3,1)	CI	SEM
Intra VMA rotation	0.00	0.06 to 0.05	1.53	1.63 to 1.43	−1.52	−1.42 to 1.62	1.53	0.983	0.985 to 0.980	0.10
Intra DM rotation	−0.20	0.11 to −0.51	8.08	8.62 to 7.54	−8.48	−7.94 to 9.02	8.28	0.625	0.671 to 0.573	2.59
Intra VMA Translation, %	0.01	0.09 to −0.07	2.21	2.35 to 2.06	−2.19	−2.04 to −2.33	2.20	0.931	0.941 to 0.920	0.29
Intra DM translation, %	−0.36	0.08 to −0.80	11.39	12.16 to 10.63	−12.11	−11.34 to −12.87	11.75	0.172	0.249 to 0.092	5.46
Inter VMA rotation	0.03	0.11 to −0.05	2.18	2.32 to 2.04	−2.11	−1.97 to −2.25	2.15	0.958	0.967 to 0.948	0.22
Inter DM rotation	−0.07	0.30 to −0.45	9.80	10.44 to 9.16	−9.95	−9.31 to −10.59	9.88	0.551	0.626 to 0.471	3.38
Inter VMA translation, %	−0.31	−0.17 to −0.46	3.58	3.84 to 3.33	−4.21	−3.96 to −4.47	3.90	0.820	0.856 to 0.778	0.84
Inter DM Translation, %	−0.62	−0.16 to −1.09	11.81	12.61 to 11.00	−13.05	−12.24 to −13.86	12.43	−0.019	0.068 to −0.096	6.92

CI, confidence interval; CR, coefficient of repeatability; DM, digitized manual vertebral motion analysis; Inter, interobserver; Intra, intraobserver; ICC, intraclass correlation coefficient; LOA, limits of agreement; SEM, standard error of measurement; VMA, computer-assisted vertebral motion analysis.

measurements (A_1–A_2, B_1–B_2, C_1–C_2) for each intervertebral level were used to establish the intraobserver agreement. The differences between each pair of observer's first round measurements (A_1–B_1, A_1–C_1, B_1–C_1) were used to establish the interobserver agreement for each measurement technique. The mean (d) and standard deviation (SD) of each set of differences were calculated. The coefficient of repeatability (CR) and upper and lower limits of agreement (LOA) were defined as follows:

$$CR = 1.96(SD)$$

$$LOA = d \pm 1.96(SD)$$

Confidence intervals were defined for the bias (d) and LOA based on the methods described by Bland and Altman [23]. List-wise deletion was employed such that only intervertebral levels with complete data across all observers and measurement techniques were included in the analysis (167 of 205). This analysis was conducted using MATLAB (MathWorks, Natick, MA, USA).

Overall intraclass correlation coefficients (ICC) and the corresponding standard error of measurement (SEM) were also calculated. Type (3,1) ICC were used according to Shrout and Fleiss and calculated using SPSS (version 19; SPSS Inc., Chicago, IL, USA) [24]. The SEM was obtained using the following formula:

$$SEM = SD\sqrt{1 - ICC}$$

Results

The average measurements and level by level contributions are demonstrated in Tables 1 and 2. Table 3 includes

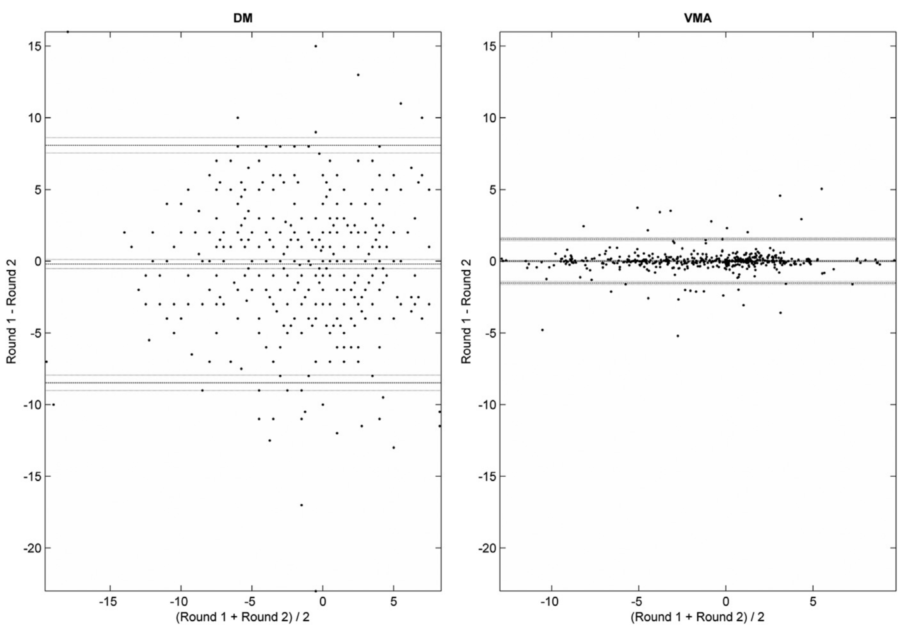

Fig. 2. Intraobserver agreement: intervertebral rotation. Bland-Altman plots of digitized manual (DM) versus computer-assisted vertebral motion analysis (VMA) measurements on axes of the same scale. Each data point is represented on the graph, along with dashed lines representing the bias, limits of agreement, and surrounding confidence intervals (thin dashed lines).

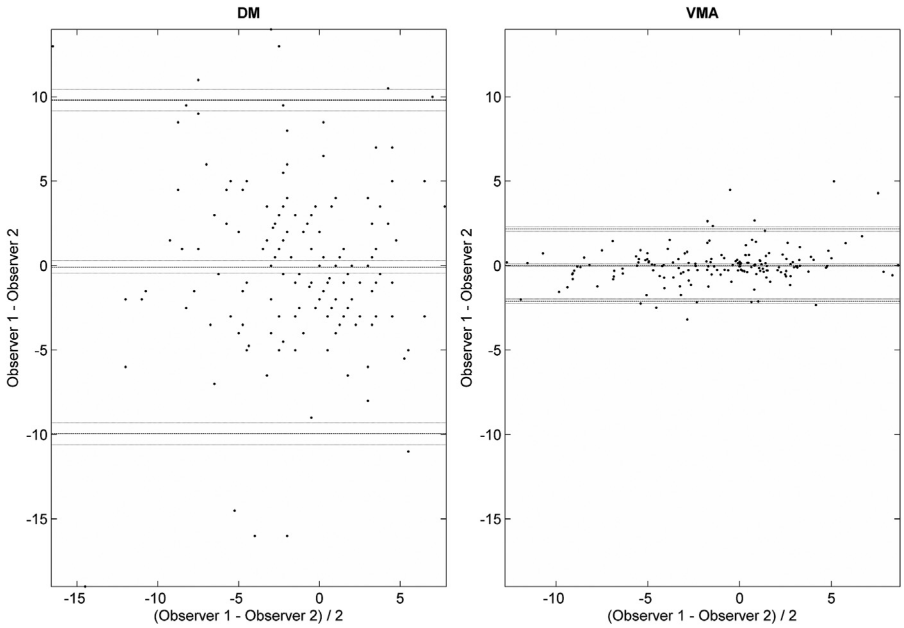

Fig. 3. Interobserver agreement: intervertebral rotation. Bland-Altman plots of digitized manual (DM) versus computer-assisted vertebral motion analysis (VMA) measurements on axes of the same scale. Each data point is represented on the graph, along with dashed lines representing the bias, limits of agreement, and surrounding confidence intervals (thin dashed lines).

the summary statistics for inter- and intraobserver agreement for the two measurement techniques. Figure 2 includes Bland-Altman plots of the intraobserver agreement for rotation measurements obtained using the manual (DM) and computer-assisted (VMA) methods on separate axes of the same scale. Each data point is represented on the graph along with dashed lines representing the bias, LOA, and confidence Intervals. Figure 3 shows the interobserver agreement on the same dataset. Analogous information is provided for the intra- and interobserver agreement for translation in Figs. 4 and 5, respectively. The non-overlapping confidence intervals for the LOA between the two measurement techniques show that significant improvements in intra- and interobserver agreement are achieved in the automated VMA technique compared with the DM technique. For intra- and interobserver rotation agreement for both techniques, the

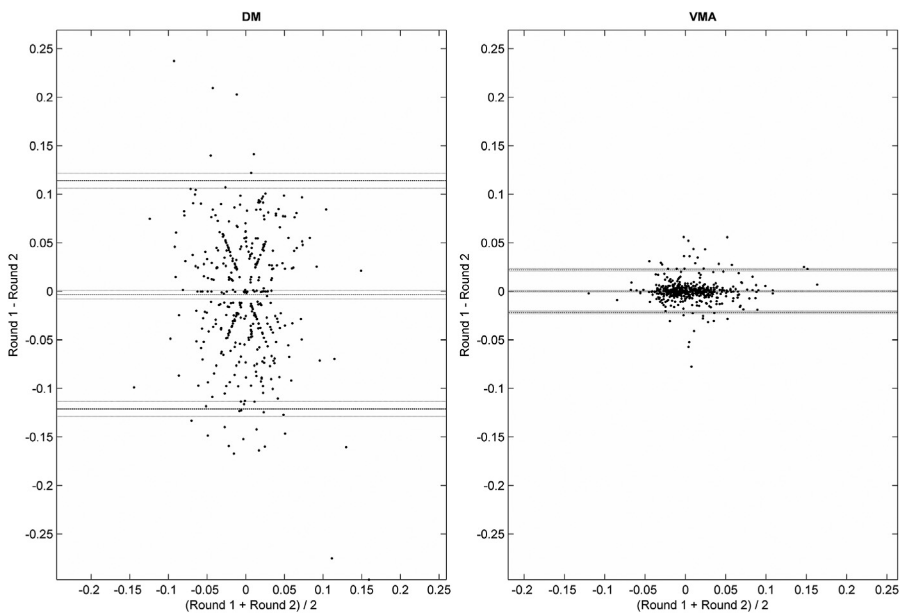

Fig. 4. Intraobserver agreement: intervertebral translation. Bland-Altman plots of digitized manual (DM) versus computer-assisted vertebral motion analysis (VMA) measurements on axes of the same scale. Each data point is represented on the graph, along with dashed lines representing the bias, limits of agreement, and surrounding confidence intervals (thin dashed lines).

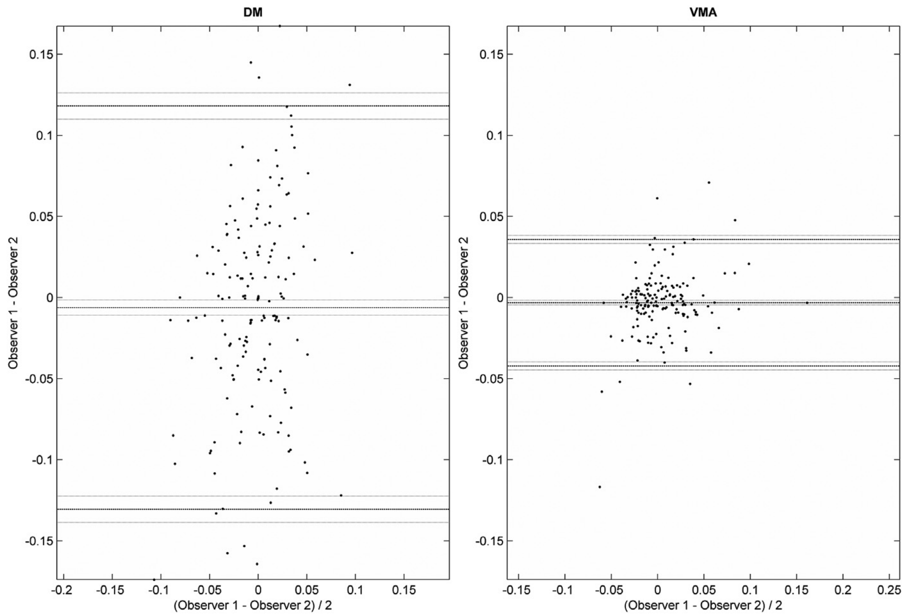

Fig. 5. Interobserver agreement–intervertebral translation. Bland-Altman plots of digitized manual (DM) versus computer-assisted vertebral motion analysis (VMA) measurements on axes of the same scale. Each data point is represented on the graph, along with dashed lines representing the bias, limits of agreement, and surrounding confidence intervals (thin dashed lines).

confidence interval for the bias includes zero, indicating that the value of the bias is not statistically significant. The same was true for intraobserver translation agreement for both techniques, but the confidence interval for interobserver translation did not include zero for either technique (Fig. 5).

Discussion

The current study demonstrates that marked improvements in intra- and interobserver precision are achieved by VMA analysis compared with a digitized manual technique. Across techniques, as expected, interobserver measurements were less reliable than intraobserver, and translation measurements were less reliable than rotation. The most reliable measurement was observed for VMA intraobserver IVR with a CR of 1.53°. This represents the inherent repeatability of the system, and suggests that a physician making repeated measurements on the same image can be 95% confident they will fall within ±1.53° of the actual value. For multiple physicians making a repeated measurement on the same image, which represents the more clinically relevant scenario, the interobserver CR for IVR was 2.15°. As expected, the least reliable measurement overall was interobserver IVT (VMA/DM CR=3.90/12.43%).

Graphical representation of both methods by expression of the LOA and surrounding CIs on adjacent vertical axes of the same scale provides visual confirmation of the improved agreement seen using automatic VMA measurements (Figs. 2–5). These plots provide a clear and intuitive interpretation of the differences between measurement methods. The marked difference between the automatic and manual CRs for all methods can also be seen in Fig. 6. Although manual measurement of DVF end range images is not standard of care, this study provides a formal reliability comparison that is not currently available in the literature. Although DVF provides significant advantages over standard radiographs, most notable of which are the ability to capture continuous motion sequences and reduce radiation exposure, the tradeoffs are also significant. The primary limitations of reduced image quality and beam parallax (the further away an object is from the conical center of a fluoroscopic beam, the greater the distortive effect of the recorded image) must be overcome. Thus, the ability of automated DVF analysis to markedly improve reliability of measurements to within the range of other available automated systems using standard radiographs should be noted.

Comparison across the literature is difficult due to the variability between methods and statistical considerations. This

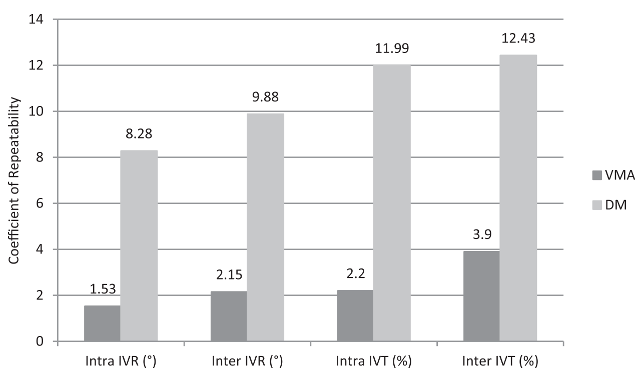

Fig. 6. Comparison of coefficients of repeatability. Computer-assisted (vertebral motion analysis [VMA]) versus digitized manual (DM) techniques. IVR, intervertebral rotation (°); IVT, intervertebral translation (% vertebral body depth).

Table 4
Comparison of reliability studies

		Methods							Statistics		
Study	Technique	Segment	Intra	Inter	IVR	IVT	X-ray	DVF	SD	SEM	ICC (type)
Current study	VMA	L1–S1	✓	✓	✓	✓		✓	✓	✓	(3,1)
Pearson, 2011 [5]	Quantitative Motion Analysis	L1–S1	✓	✓	✓	✓	✓			✓	(3,1)
Champain, 2006 [19]	SpineView	L1–S1	✓		✓	✓	✓		✓		n/a
Teyhen, 2005 [18]	DCRA	L3–S1	✓	✓	✓	✓		✓	✓	✓	(2,1)
Penning, 2005 [4]	Registration	L1–S1	✓		✓		✓		✓		n/a
Frobin, 1996 [15]	DCRA	L3–S1	✓	✓	✓	✓	✓		✓		n/a

DVF, digital videofluoroscopy; ICC, intraclass correlation coefficient; Inter, interobserver; Intra, intraobserver; IVR, intervertebral rotation; IVT, intervertebral translation; SD, standard deviation; SEM, standard error of measurement.

study focused on Bland-Altman analysis; however, ICCs and SEM were also calculated to allow for a broader comparison with other published results. Further, it has been suggested that multiple statistical methods be employed when evaluating reliability [25]. A brief comparison of select reliability studies can be seen in Table 4, along with the corresponding results in Table 5. Again, study results should be interpreted cautiously because of the different methods (range of segments, number of levels evaluated, measurement technique, etc) and statistics—for example, ICC(2,k), ICC(3,1), SEM, SD—employed. For example, ICC(2,k) coefficients are more conservative and may yield lesser values than ICC(3,1) because they incorporate both systematic and random error, where ICC(3,1) considers only the latter. Also, SEM is dependent on the ICC and will always yield values less than the standard deviation when the ICC is greater than zero.

In one reliability study evaluating computer-assisted radiographic measurements, Frobin et al. [15] reported a standard deviation of 0.99° for intraobserver IVR using a distortion compensated roentgen analysis (DCRA) technique. Teyhen et al. [18] later used a modification of this technique on DVF images and reported an error of 1.18° for intraobserver IVR. The resulting CRs are 1.94° and 2.31°, respectively, which are greater than of those observed in this study. Intraobserver IVTs were also calculated, both of which also yielded results within the range of current findings.

Only one previous study to date has evaluated and compared the reliability of automated and manual lumbar IVR and IVT measurements. Pearson et al. [5] used ICC's and SEM to demonstrate greater precision for automated Quantitative Motion Analysis measurements compared with a digitized manual technique. The ICC's reported are higher than those found in this study; however, the SEM is also greater, reflecting a larger standard deviation. This highlights the paradox often found between ICCs and SEM/SD, and the subsequent difficulty in interpreting the results.

Limitations and future work

A small but statistically significant bias was found for interobserver IVT in this study. Although this represents systematic error, it is consistent with previous findings [15,18]. Also, the corresponding confidence interval boundaries (<1.1%) are not likely to be clinically relevant.

The average magnitude of flexion and extension angles measured (4.4° and 2.0°, respectively) were smaller than others reported in the literature [4,5,15]. The same is true for IVT measurements (2.1% and 1.1%). The reduction in ROM is primarily a result of evaluating flexion and extension sequences independently. This likely increases measurement error, because the signal to noise ratio increases with decreased magnitude of motion. This is especially evident for translation, where small displacements yield the least reliable measurements.

Inter-method results (VMA vs. DM) were not reported in this study. Bland-Altman analysis is most commonly used to evaluate inter-method agreement, indicating that a new method is interchangeable with or more appropriate than the old. In the case where reliability of the old method is poor; however, little to no information can be gained [23].

Table 5
Comparison of intervertebral rotation reliability results across studies

Reliability results across studies, IVR	Intraobserver				Interobserver			
	SD	CR	ICC	SEM	SD	CR	ICC	SEM
Current study	0.78	1.53	0.983	0.10	1.09	2.15	0.958	0.22
Pearson, 2011 [5]	x	x	0.997	0.50	x	x	0.976	x
Champain, 2006 [19]	0.70–1.30	1.40–2.60	x	x	x	x	x	x
Teyhen, 2005 [18]	1.18	(2.31)	0.966–0.983	0.40–0.72	x	x	x	x
Penning, 2005 [4]	0.44–0.48	(0.88–0.96)	x	x	x	x	x	x
Frobin, 1996 [15]	0.99	(1.94)	x	x	0.9–1.60	(1.7–3.28)	x	x

CR, coefficient of repeatability; ICC, intraclass correlation coefficient; IVR, intervertebral rotation; SD, standard deviation; SEM, standard error of measurement.

Note: Results shown in parenthesis signify CR value was not listed in the respective study and calculated based on the method used in this study.

The current study focused on reliability of overall lumbar measurements made from L1–L2 to L5–S1; however, differences between levels were not analyzed. Future analysis would be of value to compare the precision of independent levels, especially in postoperative patients where instrumentation has the potential to influence measurement error. This would be of particular importance when evaluating success of fusion constructs, as well as performance of motion preservation devices.

A significant advantage of DVF is the ability to capture motion characteristics of the spine throughout an entire ROM. Further research should focus on evaluating the reliability of intervertebral motion measurements throughout the entire image sequence. This will require the use of more complex statistical analyses than are currently in use for evaluating the repeatability and reproducibility of clinical measurement systems. These techniques must consider the time series of intervertebral kinematic measurements as opposed to the discrete, end-range measurements presented here and elsewhere in the literature.

Conclusion

This study has demonstrated that reliability of intervertebral measurements by the VMA system is markedly improved when compared with a digitized manual technique. Rotational and translational measurements obtained are reproducible with average interobserver CRs of less than 2.2° and 4.0%, respectively. These results particularly highlight the ability of automated tracking algorithms to overcome challenges specific to DVF, because the errors obtained are comparable with those of previously documented plane film analysis techniques.

Acknowledgments

The authors thank Richard Prostko, MD, Hyun Bae, MD, Antonio Castellvi, MD, Reginald Davis, MD, Dan Peterson, MD, and Thomas Day, MD, for patient enrollment; Kamran Kahn, DO, Harvinder Bedi, MD, Paul Kim, MD, Alexander Yu, MD, Diana Jho, MD, and Abraham Sabersky, MD, for image analysis and measurements; and Adam Deitz, Ken Walker, Kristin Madl, Jeany Choi, and William Malpica for data processing and support.

References

[1] Deyo RA, McNiesh LM, Cone RO 3rd. Observer variability in the interpretation of lumbar spine radiographs. Arthritis Rheum 1985;28:1066–70.

[2] Shaffer WO, Spratt KF, Weinstein J, et al. 1990 Volvo Award in clinical sciences. The consistency and accuracy of roentgenograms for measuring sagittal translation in the lumbar vertebral motion segment. An experimental model. Spine 1990;15:741–50.

[3] Cakir B, Richter M, Kafer W, et al. Evaluation of lumbar spine motion with dynamic X-ray–a reliability analysis. Spine 2006;31:1258–64.

[4] Penning L, Irwan R, Oudkerk M. Measurement of angular and linear segmental lumbar spine flexion-extension motion by means of image registration. Eur Spine J 2005;14:163–70.

[5] Pearson AM, Spratt KF, Genuario J, et al. Precision of lumbar intervertebral measurements: does a computer-assisted technique improve reliability? Spine 2011;36:572–80.

[6] Frymoyer JW, Hanley EN Jr, Howe J, et al. A comparison of radiographic findings in fusion and nonfusion patients ten or more years following lumbar disc surgery. Spine 1979;4:435–40.

[7] Larsen JM, Capen DA. Pseudarthrosis of the lumbar spine. J Am Acad Orthop Surg 1997;5:153–62.

[8] Kuslich SD, Ulstrom CL, Griffith SL, et al. The Bagby and Kuslich method of lumbar interbody fusion. History, techniques, and 2-year follow-up results of a United States prospective, multicenter trial. Spine 1998;23:1267–78.

[9] Huang RC, Girardi FP, Cammisa FP Jr, et al. Long-term flexion-extension range of motion of the prodisc total disc replacement. J Spinal Disord Tech 2003;16:435–40.

[10] McAfee PC, Cunningham B, Holsapple G, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part II: evaluation of radiographic outcomes and correlation of surgical technique accuracy with clinical outcomes. Spine 2005;30:1576–83.

[11] American Medical Association. Guides to the evaluation of permanent impairment. 5th ed. Chicago, IL: American Medical Association, 2000: 379.

[12] Dupuis PR, Yong-Hing K, Cassidy JD, Kirkaldy-Willis WH. Radiologic diagnosis of degenerative lumbar spinal instability. Spine 1985;10:262–76.

[13] Hanley EN Jr. The indications for lumbar spinal fusion with and without instrumentation. Spine 1995;20(24 Suppl):143S–53S.

[14] Brinckmann P, Frobin W, Biggemann M, et al. Quantification of overload injuries to thoracolumbar vertebrae and discs in persons exposed to heavy physical exertions or vibration at the work-place: the shape of vertebrae and intervertebral discs — study of a young, healthy population and a middle-aged control group. Clin Biomech 1994;9(1 Suppl):S3–83.

[15] Frobin W, Brinckmann P, Leivseth G, et al. Precision measurement of segmental motion from flexion-extension radiographs of the lumbar spine. Clin Biomech 1996;11:457–65.

[16] Muggleton JM, Allen R. Automatic location of vertebrae in digitized videofluoroscopic images of the lumbar spine. Med Eng Phys 1997;19:77–89.

[17] Zhao K, Yang C, Zhao C, An K-N. Assessment of non-invasive intervertebral motion measurements in the lumbar spine. J Biomech 2005;38:1943–6.

[18] Teyhen DS, Flynn TW, Bovik AC, Abraham LD. A new technique for digital fluoroscopic video assessment of sagittal plane lumbar spine motion. Spine 2005;30:E406–13.

[19] Champain S, Benchikh K, Nogier A, et al. Validation of new clinical quantitative analysis software applicable in spine orthopaedic studies. Eur Spine J 2006;15:982–91.

[20] Schulze M, Trautwein F, Vordemvenne T, et al. A method to perform spinal motion analysis from functional X-ray images. J Biomech 2011;44:1740–6.

[21] Breen AC, Muggleton JM, Mellor FE. An objective spinal motion imaging assessment (OSMIA): reliability, accuracy and exposure data. BMC Musculoskelet Disord 2006;7:1.

[22] Stokes IA, Frymoyer JW. Segmental motion and instability. Spine 1987;12:688–91.

[23] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.

[24] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420–8.

[25] Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217–38.