Reprinted from SPINE, Vol. 24, No. 20, October 15, 1999

The Quantification of Low Back Disorder Using Motion Measures

Methodology and Validation

William S. Marras, PhD, Sue A. Ferguson, PhD, Purnendu Gupta, MD, Smarajit Bose, PhD, Mohamad Parnianpour, PhD, Jung-Yong Kim, PhD, and Robert R. Crowell, MD

Study Design. Trunk angular motion features were used as a means to quantify the extent of a low back disorder in healthy people and patients with chronic low back disorders.

Objective. To refine and validate a previously reported means of quantifying the extent of a low back disorder.

Summary and Background. Many assessment tools of low back disorder are subjective. A quantitative assessment tool would facilitate the tracking of the recovery and help document the appropriateness of treatments.

Methods. The trunk motion characteristics of 374 healthy people and 335 patients with chronic low back disorders of varying severity were documented as they flexed and extended their trunks in five different planes of motion. The trunk motion features were normalized as a function of age and gender. Four classification techniques were used to assess the ability of the quantitative motion measure to identify those with and without low back disorders. In addition, 31 patients were observed longitudinally to determine whether the motion measures agreed with observed changes in back pain symptoms.

Results. The quantitative trunk motion measure distinguished between people with low back disorders and healthy people between 88% and 94% of the time, depending on which classification system was used. Sensitivity and specificity varied between 83% and 97%. The quantitative measure also showed promise as a means to distinguish between muscle-based and structure-based low back disorders. Prospective findings indicated that the quantification system agreed well with clinical observations of progress.

Conclusions. The quantification of trunk motion can serve as a measure of the extent of a low back disorder. When considered along with other clinical information, the ability to assess and treat low back disorders is enhanced.
[Key words: low back pain, quantification, dynamic functional performance] **Spine 1999;24:2091–2100**

Low back disorders (LBDs) continue to represent the most common form of work-related musculoskeletal disorder[5] resulting in substantial costs to society.[6,15] In spite of this knowledge, some have observed that there has been no progress in the control of LBDs.[6] A barrier to progress has been the inability to manage LBDs consistently. To facilitate LBD treatment strategies, it is desirable to quantify and track changes in the extent of LBD.[14] Severity measures have been recognized as an essential element of medical assessment.[50] Specifically, quantification of function has been regarded as a key to functional restoration.[47] However, an accurate assessment can be problematic. Pathoanatomic diagnosis is rare,[11] and the assessment of LBD has been subjective at best.

Quantitative assessment of LBD is important for several reasons. First, without a quantitative measure of low back status, identification of back pain and disorder is difficult. In the absence of a quantitative measure, it is difficult to separate normal variability in pain from a disorder. Second, without a quantitative measure of LBD, treatment management becomes subjective. Patients have different pain thresholds, and it can become difficult to distinguish different pain tolerances from different levels of impairment. Finally, quantitative measures may provide a benchmark of when the patient is ready to return to a job without increasing the risk of exacerbating the disorder. This is important, because exacerbation of the LBD could also contribute to chronicity and, ultimately, to disability. Thus, society would benefit greatly from a quantitative measure of LBD.

Quantifying the extent of an LBD is necessary as a measure of physical impairment. Physical impairment should not be confused with other LBD definitions. It is determined through an objective assessment of structural limitations, and its determination is solely a medical responsibility. Physical impairment relates to a pathologic or anatomic loss (abnormality of structure) or a physiologic loss or limitation leading to loss of ability (functional impairment).[3,50] Disability, however, is assessed based on a patient's subjective report. The United States Bureau of Disability Insurance[49] specifies that loss or limitation evaluation be objective and "demonstrable by medically acceptable clinical and laboratory diagnostic techniques."

Quantification of low back impairment has traditionally been extremely difficult and elusive. Currently, impairment ratings of LBD vary by as much as 70%.[13] Spratt et al[46] estimated that a precise diagnosis is unknown in 80% to 90% of disabling LBDs, emphasizing the need for more quantitative measures. Traditionally, attempts to judge impairments have attempted to identify a pathoanatomic source of the LBD. Imaging techniques such as computed tomographic (CT) scans, mag-

From the Biodynamics Laboratory, The Ohio State University, Columbus, Ohio.
Funding for this project was provided by the National Institute for Disability and Rehabilitation Research (NIDRR).
Acknowledgment date: August 25, 1998.
First revision date: November 30, 1998.
Acceptance date: February 18, 1999.
Device status category: 11.

Table 1. Summary of Literature Review

Methodology	Authors	No. of Subjects	Objective of Study	Methods	Statistical Methods and Major Findings	Sensitivity and Specificity and Error Rate
Temporal aspects/ flexion–relaxation	Grabner et al. 1992	6 Controls 7 low back pain	Evaluate paraspinal muscles in controls and LBP group	Surface EMG erector spinae temporal and amplitude	The time-normalized bilateral paraspinal envelope was significantly different.	Not reported
	Ahern et al. 1988	40 LBP patients and 40 controls	Examine EMG patterns in CLBP patients and controls.	Activity levels of surface EMG measured	Multivariate discriminant function model of EMG activity	Sensitivity 84.6% Specificity 87.5% Error 14%
	Triano and Schultz 1987	41 LBP patients and 7 controls	Measure ROM, strength ratio of ext/flex, and flexion–relaxation phenomenon	Surface EMG measured on erector spinae and rectus abdominus	Two-way ANOVA was used. Flexion-relaxation was absent in patients.	Not reported
	Paquet et al 1994	10 Normals and 10 LBP patients	Compare hip–spine motion in controls and patients	Performance measured using EMG and electrogoniometers	Relaxation index was significantly higher in patients than in controls.	Not reported
Fatigue analysis or frequency analysis	Peach et al 1998	18 Normals 21 LBP patients	Develop model to classify normal and patients	Surface EMG analysis included median frequency, slope	Logistic regression model of initial median frequency and slope from the dominant multifidus was found	Test: sensit. = 100%, specif. = 75%
	Roy et al 1995	Training: 28 LBP, 42 Controls, Test: 57 LBP, 6 Controls	Distinguish muscle impairment between patients and controls	Surface EMG measuring initial median frequency and median frequency slope	Discriminant analysis model of IMF and MF slope	Training: sensit. 85% specif. 86% Error 14% Test: sensit. 88% specif. 100% Error 6%
	Klein et al 1991	17 Non-LBP rowers and 8 LBP rowers	Difference between athletes and LBP athletes	Surface EMG of low back musculature	Discriminant function model using percentage median frequency recovery.	Sensitivity 66% Specificity 71% Error rate 31%
	Mayer et al 1989	11 healthy controls, 10 LBP patients	Spectral EMG differences in patients and controls	Surface EMG measuring spectral parameters	T test showed significant differences on spectral slopes between controls and patients.	Sensitivity 40% No specificity reported
	Biedermann et al 1991	22 Normals and 27 patients (avoider 9 and confronter 15)	The objective was to systematically evaluate spectral parameters	Surface EMG of the iliocostalis muscle	Discriminant analysis model of spectral parameters identified 88.9% of avoiders, 67% of confronter and 59% of controls	Sensitivity 78% Specificity 59% Error rate 31%
EMG activity level	Sherman 1985	15 Controls, 28 history of LBP, 83 with LBP symptoms.	The objective was to distinguish among the groups	Measure surface EMG of paraspinal muscles in different postures	No significant difference between controls and history of symptom group. Current LBP subjects had EMG activity above controls.	Not reported
	Cooper et al 1993	28 Normals, 20 Patients	Examined paraspinal muscle function in Controls and LBP	EMG on paraspinal muscles	IEMG was significantly greater in patients compared to controls.	Not reported
	Alexiev 1994	40 Normal, 40 LBP patients	Compare strength and EMG in controls and LBP patients	Measure trunk position with goniometer and EMG activity of erector spinae	Controls were significantly stronger than patients but did not have significantly higher iEMG.	Not reported
	Robinson et al 1992	16 CLBP patients and 12 Controls	Describe strength and iEMG in controls and CLBP	Surface IEMG of paraspinal muscles	The groups showed significantly different IEMG fatigue slopes.	Not reported
Strength testing	Mandell et al 1993	21 Controls US postal workers 59 LBP patients	Distinguish controls and LBP patients with functional perform. measures.	Evaluated ROM, strength and cardiovascular fitness	Isometric and isokinetic peak force and torque tests failed to show significant differences between low back pain and workers.	Not reported
	Kishino et al 1985	65 Normals and 68 patients	Compare controls to chronic LBP patients	Isokinetic strength	Paired *t* tests showed that patients lifted significantly less weight compared to controls.	Not reported
	Kumar et al 1995	73 controls, 10 LBP	Develop database of controls	Flexion extension and lateral flexion strength	The patient isokinetic strength was 9% to 40% less than the reference controls group.	Not reported
	Mayer et al 1984	12 Control 38 Chronic LBP	Measure spinal ROM in controls and patients.	Evaluated ROM using inclinometers and x-ray	Descriptives for two groups but not statistics to compare controls and patients	Not reported
	McIntyre et al 1991	32 Normals, 23 Patients	Compare controls to LBP patients.	Measure strength ROM, velocity and oscillation angle	Average velocity was significantly different between controls and patients at $P < 0.001$.	Not reported

Table 1. (Continued)

Methodology	Authors	No. of Subjects	Objective of Study	Methods	Statistical Methods and Major Findings	Sensitivity and Specificity and Error Rate
Strength testing	Hultman et al 1993	36 Normals 91 Intermittent LBP, 21 CLBP	Compare groups in strength, endurance, and body composition	Trunk strength measures at 30°/s Calipers measure body composition	Significant difference in isometric endurance between normals and intermittent LBP as well as between CLBP and intermittent LBP	Not reported
	Newton et al 1993	70 Normals and 120 CLBP	Standardize testing methods with controls and LBP group	Measure strength, ROM, fatigue and endurance	ICCs on controls were above 0.80. Differences in sex but not age or body weight.	Sensitivity: 85–90% Specificity 77–80% Error 15–19%
	Lee et al 1995	37 Normals 61 LBD	Compare strength in trunk and knee between group	Measure isokinetic strength at 120 deg/sec knee strength at 126°/s	Tests showed that the LBP group had significantly lower strength in both the trunk and knee compared to the normal group	Not reported
	Ito et al 1996	90 Normals 100 CLBP	Compare normals and CLBP group	Endurance was measured in time person could hold position	Tests showed that the normals group had longer endurance time than CLBP patients.	Not reported
	Brady et al 1994	191 Chronic LBP Compared to normative data	Evaluate isokinetic trunk strength before and after functional restoration	Trunk strength measured at 60°/s and 150°/s	Studentized t-test showed significant improvement. The % normal analysis indicated residual strength decrements.	Not reported
Motion measures	Mellin et al 1989	143 Normals 157 inpatients 156 Outpatients	Compare modes of treatment	Inclinometer measures of lumbar flexion and extension, hip mobility.	No difference in treatment groups. Hip and lumbar motion improvement correlates with subjective progress more than trunk strength.	Not reported
	Jayarman et al 1994	12 Healthy 10 LBP	Compare motions between normal and painful spines	Video analysis was performed.	Wilcoxon-Mann-Whitney test showed significant difference between control and patients for position and velocity depending on segment level.	Not reported
	Marras et al 1995	339 Normals 171 Patients	Compare motions between normals and patients	Lumbar motion monitor measuring pos., velocity, acc.	Discriminant function, splines and CART were used to classify control and patients with 95% to 89% accuracy.	Error rate = 4.9–11.8% depending on statistical method
	McClure et al 1997	12 Normals 12 Prior LBP	Compare lumbar and hip motion pattern in those with & without history of LBP	3-D optoelectric motion analyzer	Test showed the group with history of LBP had greater lumbar motion and velocity during the initial phase of extension.	Not reported

netic resonance imaging (MRI), and myelograms are used to assist in the identification of the structure that has been compromised. However, more than 85% of LBDs do not have a pathoanatomic diagnosis.[11] Current imaging techniques often are not sensitive enough to observe anatomic anomalies throughout the functional range.

Given these difficulties, recent attempts to develop a measurement system for LBD have centered around functional measures of impairment. Several approaches or methods have been identified in the literature. These include the use of electromyogram to observe flexion–relaxation and/or temporal features of the torso muscles, observation of muscle fatigue by electromyogram spectral analysis, observation of electromyogram activity level, determination of trunk strength, and documentation of trunk motion components. Table 1 summarizes the features of the more notable studies. Several limitations are associated with many of these studies. First, many of of them involved small samples, which limits the application of results to the general population. Second, examination of the sensitivity/specificity column in Table 1 indicates that the issues of classification error has been considered in few of these studies. Furthermore, few of the investigators have been able to identify patients with LBD from a sufficiently large (normally distributed) population of healthy subjects and subjects with LBDs. Finally, few investigators have attempted to validate their findings using multiple robust evaluation techniques.

An alternative approach to assessing the extent of an LBD has been suggested by Marras and Wongsam[27] and Marras et al.[26,29–31] This approach documents the symmetric and asymmetric back motion characteristics of a patient and compares these motion characteristics with those of a healthy, unimpaired subject group adjusted for age and gender. Patients are tested in different torso asymmetries so that different combinations of the trunk's muscles must be recruited to flex and extend the trunk. The motion profile observed during repeated flexion and extension of the torso at different trunk asymmetries is believed to be a reflection of the trunk's musculoskeletal central control program, often called the "central set."[17] In unimpaired subjects, we believe that this control program is well developed during the subject's lifetime. However, in an injured patient, it is thought that this

musculoskeletal control program must be adjusted to compensate for limitations related to muscle functions, structural restrictions, and guarding behavior. Marras and Wongsam[27] were able to show that the best discriminators between healthy people and patients with LBD involved more dynamic trunk motion indicators (velocity). Later, trunk velocity and acceleration motion characteristics as well as expected variability for a large group of healthy subjects was described.[31] The results indicated that age and gender could influence trunk motion characteristics and that motion characteristics could be normalized for age and gender. In more recent studies,[29,30] we compared the activities of these healthy subjects with those of patients with LBD and concluded that trunk motion measures are repeatable and that only dynamic motion characteristics (not ROM) vary as a function of LBD. In the later studies, we attempted to distinguish patients with LBD from healthy subjects and to classify the type of LBD into 1 of 10 categories, based solely on trunk motion characteristics. Eight motion variables were used to distinguish between healthy subjects and patients with LBD, with very good resolution. This motion model serves as the basis for a quantification system for LBD. A second eight-variable model was used to further classify patients with LBD into 1 of 10 diagnostic categories. However, only one of the four classification methods indicated that this second model would be capable of classifying subjects into the 10 diagnostic categories accurately.

Using this body of knowledge as a base, the objective of the current study was to refine and validate further the LBD motion quantification model. This was accomplished by determining how well the quantification tool distinguished in large group of subjects between healthy people and people with LBD. In addition, the goal of this study was to explore the potential for use of the quantification technique to assess the nature of LBDs.

■ Methods

Approach. Currently, LBD progress is assessed primarily by subjective reports. Therefore, no quantitative continuous benchmarks or gold standards are available with which to judge trunk motion model performance. The approach taken here was to assess a group of patients with varying degrees of LBD to determine whether a combination of trunk motion measures (motion model) could be used to distinguish them from a group of healthy participants. Because different participants had varying degrees of LBD and the healthy group had different degrees of symptom-free qualities, it was expected that this study would provide a reasonable test of the continuous nature or sensitivity of the motion measure model.

Subjects. Two groups of participants, healthy subjects and patients with LBD, were recruited. Three hundred and seventy-four participants were recruited who had no history of low back pain and were considered the healthy group. This group consisted of the original healthy group reported earlier[31] plus an additional 23 subjects used to attain better balance in the representation of the various age groups. Healthy participants ranged in age from 20 to 70 years (mean, 38.9 ± 14.9 years). A minimum of 25 men and women were recruited as control subjects for each decade of age in the database. Three hundred thirty-five patients with low back pain represented the LBD group (average age, 42.8 ± 13.4 years). The patients were recruited from secondary referral practices and had had symptoms for at least 6 months. The patients with LBD were distributed among 10 symptom or diagnostic categories as defined by the Quebec Task Force Study.[30] There were 35 patients in the local low back pain category (Quebec 1) with a mean pain level of 3.6 ± 2.5. Forty-eight patients were classified with spinal stenosis and had an average pain level of 5.6 ± 2.7. Further details of pain levels within each diagnostic category are available from the authors. The average height in the healthy group was 170.9 ± 14.3 cm and in the LBD group was 168.5 ± 27.2 cm. There was no significant difference in anthropometric data between patients and healthy subjects.

Experimental Design. The method used to assess the LBD quantification measure was the same as that described by Marras et al.[30] The protocol required participants to flex and extend the trunk in symmetric and asymmetric postures while the motion components of the torso during these tasks were observed.

Two validation methods and four classification schemes were used to evaluate whether the motion model could be used to distinguish between a large database of healthy participants and patients with LBD. The first validation method (Method I) involved a cross-validation technique. In this technique, the model coefficients used for classification were developed with part of the data set and tested against the remaining data set. The process was repeated using different portions of the data set. This provided an estimate of overall model robustness.[41] The second method (Method II) of validation consisted of an independent test of the model in which part of the database was used as a training set to develop the coefficients for the quantitative model. The remaining data set was used to independently evaluate model fidelity independently. Method II provided a strong independent measure of model validity.

The four classification schemes used in the analysis consisted of the same four schemes used to analyze the original data set[30]: 1) discriminant function analysis, 2) classification using regression trees (CART), 3) classification using splines (CUS), and 4) modified classification using splines (MCUS).[8] Four classification schemes were used to assess the degree of convergence of the methods.

Finally, a prospective observation of the measures was documented. Longitudinal observations were performed in a subset of the subjects. The goal was to evaluate whether the motion measures varied, along with other more subjective LBD measures as patients progressed through their LBD overtime.

Experimental Conditions. The experimental tasks consisted of flexion–extension exertions performed while the trunk was held within one of five asymmetric planes of motion (Figure 1). In addition, twisting range of motion and the ability to perform the asymmetric tasks were used as dependent measures. These asymmetric planes of motion consisted of 1) a sagittally symmetric plane (0° of twist), 2) a 15° twist clockwise (CW) in the transverse plane, 3) a 15° counterclockwise (CCW) twist, 4) a 30° CW twist, and 5) a 30° CCW twist. The first task was the sagittally symmetric condition, followed by the two 15° conditions and then the two 30° conditions. The CW and CCW

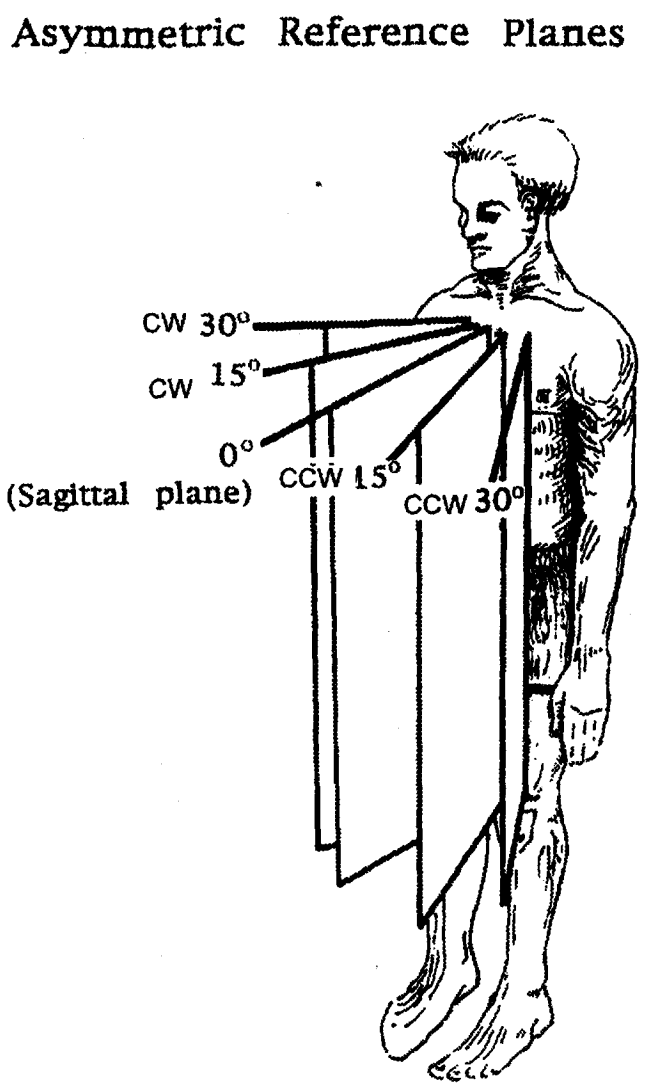

Figure 1. Asymmetric twisting position. cw, clockwise; ccw, counterclockwise.

conditions were counterbalanced at both 15° and 30°. Subjects were not always able to perform the 15° and 30° tasks, and this capability served as the ability variable. Excellent repeatability and reliability of these measures has been reported.[30]

Motion Measures. The 52 motion measures representing the trunk motion signatures described in Table 2 served as dependent variables in this study. Biomechanically, trunk motion performance was expected to decrease as tasks became more asymmetric because smaller oblique muscles are recruited and are necessary for motor control during these tasks. The theory suggests that those with healthy low backs would have a different muscle recruitment and coactivity pattern resulting in a different motion signature than those with LBD. These motion measures were then used to distinguish between healthy and impaired performance and as benchmarks for the severity of injury.

Apparatus. The lumbar motion monitor (LMM) was used to measure trunk motion performance in all three planes of the body. The LMM is an exoskeleton of the spine that is attached to the hips and shoulder using harnesses made of orthoplast (Figure 2). These harnesses serve as a stable attachment or anchoring point for the LMM in describing the motion of the spine. Because the harnesses embody the thorax and pelvis, they permit the LMM to bend between these two attachment points and follow the motion of the thoracolumbar spine, thus reflecting thoracolumbar trunk position changes. The LMM signal is collected at 60 Hz and transmitted either by hardwire or telemetry to an analog-to-digital converter on a portable computer and stored for further analysis. The LMM has been validated quantifying position, velocity, and acceleration in all three planes of the body.[28]

The experimental task required the subject to maintain the twisting position (during the flexion and extension exertion) within a tolerance of ±2° from the desired plane of motion (task asymmetry). The transverse plane position signal from the LMM was fed to a comparator circuit and displayed on an oscilloscope. This provided both a visual and auditory system of feedback to the participant to control the twisting position during the experimental task.

Table 2. Fifty-Two Motion Variables Evaluated

	Asymmetry				
Motion Measures Observed	0°	15° cw	15° ccw	30° cw	30° ccw
Sagittal plane ROM	x*	x	x	x	x
Sagittal plane peak flexion velocity	x	x	x	x	x
Sagittal plane peak extension velocity	x*	x	x	x	x
Sagittal plane peak flexion acceleration	x	x	x	x	x
Sagittal plane peak extension acceleration	x*	x	x	x	x
Frontal plane ROM	x*	x	x	x	x
Frontal plane peak right side velocity	x	x	x	x	x
Frontal plane peak left side velocity	x	x	x	x	x
Frontal plane peak right side acceleration	x	x	x	x	x
Frontal plane peak left side acceleration	x	x	x	x	x
The ability to successfully complete the task in each of the five asymmetric positions (experimental tasks)				x*	
Twisting ROM capability (maximum cw rotation minus maximum ccw rotation)				x*	

cw = clockwise asymmetry; ccw = counterclockwise asymmetry; ROM = range of motion.
* Used in the final multivariate model.

Figure 2. Lumbar motion monitor on a subject.

Procedure. Subjects were instructed to cross their arms in front of them, stand with their feet shoulder-width apart, and flex and extend their trunk as fast as they could comfortably while maintaining a twisting position within the control zone on the display. If the twisting position indicator in the display went outside the target zone, the trial was repeated.

Data Analysis. Custom software converted the LMM signal into trunk position, velocity, and acceleration. The first flexion–extension cycle was discarded from the analysis. The readings from the next four flexion–extension cycles were analyzed and averaged.

Statistical Analysis. The entire database of 709 participants (374 healthy participants and 335 patients with LBD) was used to determine classification error and sensitivity and specificity for the cross-validation method of assessment. However, for the independent test set validation, the database was split into training and testing data sets. The healthy-participant training data set contained two-thirds of healthy subjects (n = 228) from the original database.[31] The LBD-patient training set consisted of the entire data set (n = 171 patients with LBD) originally described by Marras et al.[30] The independent test set consisted of the remaining one-third of the healthy group plus an additional 23 healthy participants (n = 146). Test set patients with LBD consisted of 164 patients who participated in the previous study.

LBD Quantification. Results of previous analyses have shown that trunk motions change as a function of age and gender as well and as a function of LBD.[31] The quantification method normalizes trunk motion for age and gender and thus removes variability caused by these factors, so that trunk motion performance is purely a function of LBD. This process is accomplished by dividing the value of the specific trunk motion component of a patient with LBD by the mean value of the normative group's specific trunk motion component, given the gender and decade of age, thus describing LBD as a percentage of the expected normal value.

For both the cross-validation and independent test set validation methods, the data were normalized (for age and gender) using the control group motion characteristics derived from the 228 control subjects. This population of subjects was used as the training set for the independent test set validation. Comparisons of motion characteristics between this normalization group and the entire 374 healthy participant data set was statistically nonsignificant at the 0.01 level of significance. The normalization values were essentially the same as those reported by Marras et al.[30]

One goal of this statistical analysis was to build a simpler, yet effective, quantification model compared with that reported earlier.[30] The previous model contained two terms that required continuous assessment of the data. The previous data were re-evaluated using a simpler model that excluded the two continuous terms and minor (nonsignificant) differences in model performance resulted. Thus, given the difficulty associated with assessing continuous variables, the previous model was simplified to six variables (Table 2).

Finally, a model was developed that was intended to distinguish between muscle-related problems and structure-related problems. In the model, two complex ratio variables are considered (listed in equations 1 and 2).

$$\text{SF ACC}_{(0)} \Big/ \sum \text{SF ACC}_{(0,15cw,15ccw,30cw,30ccw)} \quad (1)$$

$$\text{SE ACC}_{(15cw)} / \text{SE ACC}_{(0)} \quad (2)$$

where SF is sagittal flexion, SE is sagittal extension, ACC is acceleration, and () is asymmetry, where 0, 15cw, 15ccw, 30cw, and 30ccw are asymmetry angle and direction.

■ Results

The differences between the performances of patients and healthy subjects for the sagittal trunk motion measures are presented in Figure 3. The figure illustrates a greater difference between the two groups in velocity and acceleration measures than in ROM. Similar trends were observed in the current study as were observed in the previous study.[30] Trunk ROM, velocity, and acceleration in the sagittal plane decreased for all subjects as the test conditions became more asymmetric. Compared with the healthy group, the ability to perform the various asymmetric tasks as well as the magnitude of the performance measures were significantly reduced in the LBD group. The greatest differences between the healthy and LBD categories related to measures of the higher order derivatives of motion (*i.e,*. velocity and acceleration). Table 3 shows results of the univariate analyses of the sagittal motion measures, which demonstrate that the

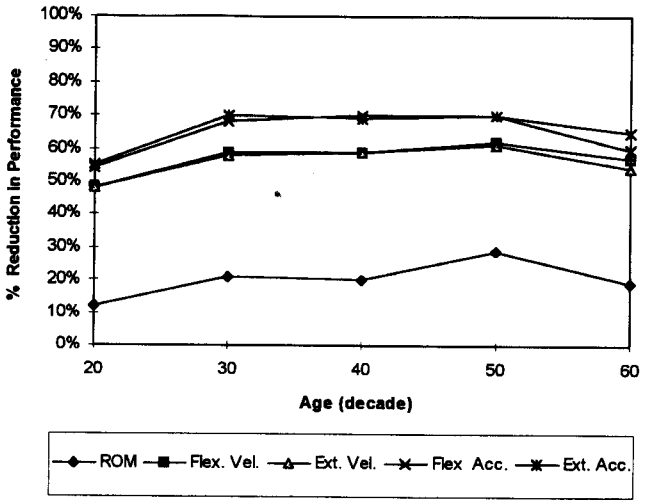

Figure 3. Percentage reduction in sagittal plane performance of control subjects compared with patients for male participants.

higher order derivatives of sagittal plane motion were better able to distinguish between healthy participants and patients with LBD. This emphasizes that velocity and acceleration measures were extremely important for distinguishing between the two groups.

Data in Table 4 indicate the ability of a combination of the trunk motion measures (multivariate analysis) to correctly classify the 709 participants in this study as patients with LBD or healthy participants. These data indicate that the cross-validation method resulted in correct classification for between 88% and 90% of the subjects among the different classification methods. Similar classification measures are presented for the independent test set validation method where the training set is evaluated independently from the test set. The classification error rates for the training set of data varied from 8.5% to 12.3%. Error rates for the test data set were even better, varying from approximately 6% to 8.7%. It is notable that these results were relatively consistent among the different types of classification and validation methods used. This consistency indicates that the model is robust and capable of distinguishing patients with LBD from unaffected subjects.

Table 4 also shows the sensitivity and specificity values for all the various validation and classification techniques. The results of this analysis indicate a consistency of sensitivity and specificity regardless of validation method or classification technique. Overall, validation and classification methods sensitivity ranged from 83% to 92%, whereas specificity ranged from 86% to 97%. All validation methods and classification techniques produced sensitivity and specificity values that were well balanced (both high in value), indicating that quantification can be performed with a high degree of accuracy, minimizing the two types of error common in diagnostic testing.

As expected, the motion model could not be used for accurate assessment of the specific category of LBD. Error rates for this classification ranged from 75% to 82%. Therefore, the potential for use of the method to classify specific LBD categories correctly was not confirmed.

The model described in equations 1 and 2 was tested for its ability to distinguish between the structural or muscular origin of LBD. Two hundred seventeen patients were correctly classified as having a muscular or structural basis for their pain. The classification results for this model indicated that 69% of the structural group and 80% of the muscular group were correctly classified. This model shows that one can distinguish well between structural and muscular problems with approximately 75% accuracy. Given that the true diagnoses for the patients within this study are often questionable, these classifications are considered a good estimate of the general nature of the LBD.

Table 3. The Error Rate for Univariate Discriminant Function Analyses Identifying Controls vs. Patients

	Discriminant Function		
	Method I	Method II	
Single Variable Control vs. Patient Models	(Cross-Validation)	Training Set	Test Set
Range of motion	0.4006	0.4196	0.3857
Sagittal flexion velocity	0.2129	0.2317	0.2128
Sagittal extension velocity	0.2133	0.2366	0.2124
Sagittal flexion acceleration	0.2277	0.2385	0.2391
Sagittal extension acceleration	0.2104	0.2194	0.2143

Table 4. Control vs. Patient Sensitivities, Specificities, and Error Rate by Statistical Method

		Statistical Method			
Data Set		Disciminant Function	CART	CUS	MCUS
Method I	Sensitivity (%)	85	90	84	
Cross-validation	Specificity (%)	95	86	94	
	Error rate	0.1002	0.1213	0.1027	
Method II	Sensitivity (%)	83	88	85	86
Training Set	Specificity (%)	93	87	96	95
	Error rate	0.1177	0.1228	0.0852	0.0877
Test Set	Sensitivity (%)	90	91	90	92
	Specificity (%)	94	92	96	97
	Error rate	0.0792	0.0870	0.0710	0.0581

Discussion

These results have confirmed that accurate objective benchmarking of an LBD can be achieved by observing motion characteristics. Using a more parsimonious motion measure model than that used in a previous study, the extent of LBD was successfully quantified. The advantage of the simpler model is that model components can be analyzed more quickly and efficiently than could the complex continuous-motion components used previously.[30]

Correct classification performance of the model in this validation study was roughly equivalent to that in our original study.[30] The model demonstrated excellent sensitivity and specificity and can be used as an accurate measure of the extent of LBD. Compared with the previous methods reporting sensitivity and specificity (Table 1), the current quantification system performs very well. Additionally, the system has been tested using a database that is more than three times larger than the next largest (nonmotion based) database reporting sensitivity and specificity. The only study involving nearly as large a sample as the current study was that by Mellin et al,[37] which involved 453 subjects. However, they could not distinguish between treatment methods (which was their goal) and did not report sensitivity or specificity.

Trunk motion assessment data were rather simple to collect, compared with the collection process in other systems that attempt to classify patients with LBD. Five of the previous studies reporting sensitivity and specificity required electromyography. This often requires disrobing and an application of strength (which may be limited by pain in some participants). The other technique reporting sensitivity and specificity required testing of torso strength.[38] Although they reported respectable sensitivity and specificity, the safety associated with strength testing is often in question. For example, Battiée et al[4] found that they injured more participants using strength testing than the number of LBDs they were able to prevent. Thus, the present method is easier to administer and has better accuracy than previous efforts.

Several features of this study indicate that the technique of quantifying LBD using motion measures is a valid approach. First, a valid method should perform well in a large number of subjects. With a large data set, questionable subject performance would be diluted by the mass of the database. This study represents the largest data set for evaluating the extent of an LBD found in the literature to date. With more than 700 participants and strong sensitivity and specificity findings, this study provides a stringent test of the method. Second, several validation and classification procedures were used, resulting in very similar error rates, sensitivities, and specificities. The convergence of the results indicates a very robust and consistent model. Third, compared with other available methods for identification of LBD, this technique not only has a lower classification error, but it has a much better balance between the strength of sensitivity and specificity. In fact, it is possible that this approach may outperform the gold standard methods that were used by the physicians for recruiting patients for this study.

A sensitivity analysis of our results was performed to determine where misclassification errors occurred. These analyses indicated that those patients misclassified as healthy had statistically ($P < 0.01$) lower pain scores by an average of 43% than the correctly classified patients. It appears that the model did not misclassify randomly but specifically, with participants whose self-reported pain was significantly less severe.

It should also be noted that the measure of LBD discussed in this six-variable model uses only a small portion of the available variables. The individual velocity and acceleration profiles observed in specific asymmetries during the flexion–extension tasks may serve as further measures of specific structural problems (*e.g.*, tenderness of a nerve root at a particular point). A documented motion deficit in one of the specific asymmetries may indicate a specific disability and may be tracked as an indication of recovery. The value of motion measures may go well beyond the general assessment reported here.

This method provides a simple, easy to use, and sensitive measure that quantifies the extent of an LBD, given the patient's age and gender. The analyses performed here evaluate the value of the motion measures in isolation from any other information. Under realistic usage conditions, it could be expected that the motion measures could be used in conjunction with other indicators of disease. For example, factors such as a medical history or straight leg raising ability may be used in conjunction with trunk motion measures to paint a more complete picture of patient status.

As expected, we were not able to validate the classification of individual diagnostic categories as hoped.[30] However, in that study results obtained with only one classification method indicated that this might be possible. No other assessment techniques have been used in such a rigorous test (Table 1). Thus, we no longer believe that motion measure models are currently capable of specific diagnoses. However, we believe that the distinction between a muscular and structural source of LBD is possible and valuable. Results in our studies have indicated that motion measure models have very different profiles when the origin of an LBD is structural rather than muscular.

Finally, the results indicate that this quantification technique can be used as a tool to track patient progress during treatments. A sensitive measure can be expected to track changes in LBD status over time. As a test of this expectation, 31 subjects were tested longitudinally on two occasions separated by several weeks. The patients showed two distinct patterns during this period. Fifteen patients reported improvement, whereas 16 reported the same or worsening pain. Figure 4 indicates that as self-reported pain improved over observation time there were

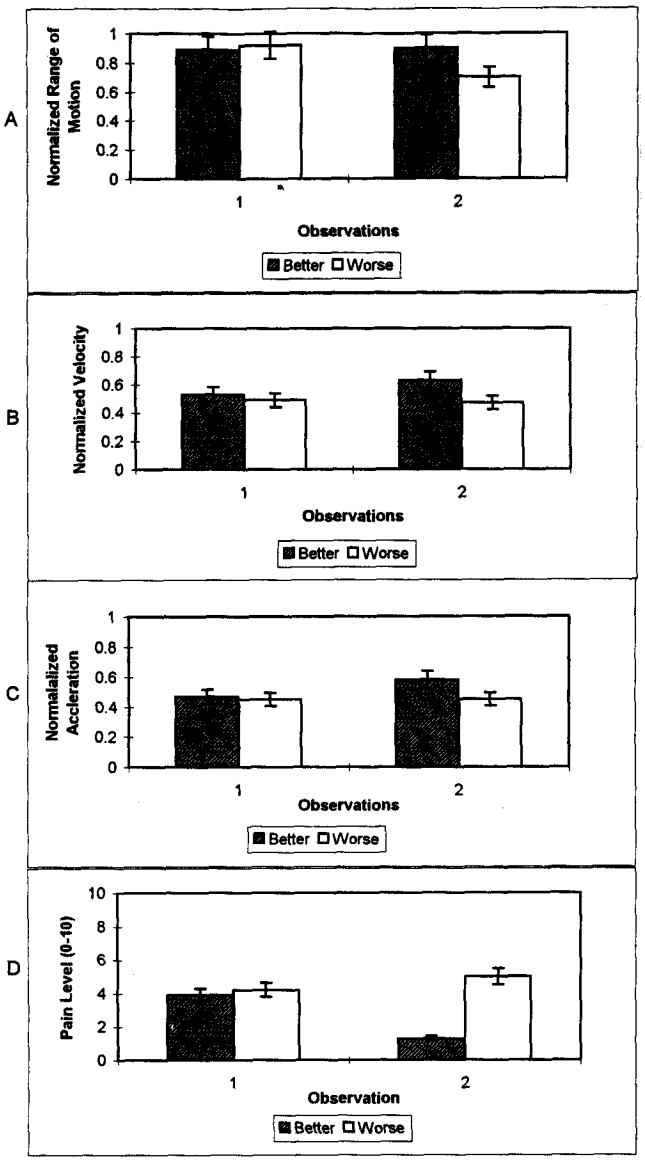

Figure 4. Range of motion **(A)**, velocity **(B)**, acceleration **(C)**, and pain **(D)** as a function of observation.

significant changes in the velocity and acceleration measures (but not in the ROM measure), contrary to those who did not report improvement. Changes in ROM were not indicative of changes in pain status, confirming earlier findings.[30] Another article[12] contains a thorough reports of how the motion measures described here are able to track medical improvement in the LBD for observation periods of 3 to 6 months.

The motion measures and the models quantifying the extent of a disorder would only be useful if the subject were producing sincere trunk motion exertions. Another article[32] contains a report of how motion measures can be used to assess sincerity of effort.

This quantitative measure of LBD can also serve as a means to track factors affecting LBD in future studies. Now that we have established a means to assess the extent of an LBD, this benchmark can be used as a means to assess and compare the effectiveness of various treatments. In addition, the measure can be used to determine when a patient may be a candidate for extreme treatments, such as surgery, or when the patient is ready to return to the workplace without exacerbation of the problem. Finally, this LBD quantification can also be used in conjunction with workplace quantification measures so that the matching of the workplace to the person's capabilities can be more scientifically approached.

Acknowledgments

The authors thank Gary Rea, MD, PhD, Deborah Burr, PhD, Sheldon Simon, MD, and Karen E. K. Lewis, MS, for their contributions to this study.

References

1. Ahern D, Follick M, Council J. Comparison of lumbar paravertebral EMG patterns in chronic low back pain patients and non-patient controls. Pain 1988; 34:153–60.
2. Alexiev A. Some differences of the electromyographic erector spinae activity between normal subjects and low back pain patients during the generation of isometric trunk torque. Electromyogr Clin Neurophysiol 1994;34:495–9.
3. American Medical Association. A guide to the evaluation of permanent impairment of the extremities and back. JAMA 1958;166(Suppl):1–122.
4. Battiée M, Bigos S, Fisher L, Hansson T, Jones M, Wortley M. Isometric lifting strength as a predictor of industrial low back pain reports. Spine 1989;14: 851–6.
5. Bernard B. Musculoskeletal Disorders and Workplace Factors: A Critical Review of Epidemiologic Evidence for Work-Related Musculoskeletal Disorders of the Neck, Upper Extremity, and Low Back. U. S. Department of Health and Human Services (NIOSH) Publication No. 1997;97–141.
6. Bigos S, Battiée M. The impact of spinal disorders in industry. In: The Adult Spine: Principles and Practice. 2nd ed. Frymoyer JW, ed., Philadelphia: Lippincott–Raven, 1997:151–61.
7. Biedermann H, Shanks G, Forrest W, Inglis J. Power spectrum analyses of electromyographic activity discriminators in the differential assessment of patients with chronic low back pain. Spine 1991;16:1179–84.
8. Bose S. A method for estimating nonlinear class boundaries in the classification problem and comparison with other methods. Department of Statistics, University of California, Berkley, 1992. PhD dissertation.
9. Brady S, Mayer T, Gatchel R. Physical progress and residual impairment quantification after functional restoration. II: Isokinetic strength. Spine 1994;19: 395–400.
10. Cooper R, Stokes M, Sweet C, Taylor R, Jayson M. Increased central drive during fatiguing contractions of the paraspinal muscles in patients with chronic low back pain. Spine 1993;18:610–4.
11. Dillane J, Fry J, Kalton G. Acute back syndrome: A study from general practice. BMJ 1966;2:82–4.
12. Ferguson S, Marras W, Gupta P. Longitudinal observations of the natural course of low back pain recovery. Spine (in press).
13. Frymoyer J. A practical guide to current United States impairment rating: A critical analysis. In: Frymoyer JW, ed. Adult Spine: Principles and Practice. New York: Raven Press, 1991:Chap 11.
14. Frymoyer J. Quality: An international challenge to the diagnosis and treatment of disorders of the lumbar spine. Spine 1993;18:2147–52.
15. Frymoyer J, Durett C. The economics of spinal disorders. In: Frymoyer JW. The Adult Spine: Principles and Practice. 2nd ed. Philadelphia: Lippincott–Raven, 1997:143–50.
16. Grabiner M, Koh T, Ghazawi A. Decoupling of bilateral paraspinal excitation in subjects with low back pain. Spine 1992;17:1219–23.
17. Horak F, Diener H. Cerebellar control of postural scaling and central set in stance. J Neurophysiol 1994;72:479–93.
18. Hultman G, Nordin M, Saraste H, Ohlsen H. Body composition, endurance, strength, cross-sectional area, and density of MM erector spinae in men with and without low back pain. J Spinal Disord 1993;6:114–23.
19. Ito T, Shirado O, Suzuki H, Takahashi M, Kaneda K, Strax T. Lumbar trunk muscle endurance testing: An inexpensive alternative to a machine for evaluation. Arch Phys Med Rehabil 1996;77:75–9.
20. Jayaraman G, Nazre A, McCann V, Redford J. A computerized technique for analyzing lateral bending behavior of subjects with normal and impaired lumbar spine. Spine 1994;19:824–32.

21. Kishino N, Mayer T, Gatchel R, et al. Quantification of lumbar function. Part 4: Isometric and isokinetic lifting simulation in normal subjects and low-back dysfunction patients. Spine 1985;10:921–7.
22. Klein A, Synder–Mackler L, Roy S, Deluca C. Comparison of spinal mobility and isometric trunk extensor forces with electromyographic spectral analysis in identifying low back pain. Phys Ther 1991;71:445–54.
23. Kumar S, Dufresne R, Van Schoor T. Human trunk strength profile in flexion and extension. Spine 1995;20:160–8.
24. Lee J, Ooi Y, Nakamura K. Measurement of muscle strength of the trunk and lower extremities in subjects with history of low back pain. Spine 1995;20: 1994–6.
25. Mandell P, Weitz E, Berstein J, et al. Isokinetic trunk strength and lifting strength measure differences and similarities between low-back-injured and non-injured workers. Spine 1993;18:2491–501.
26. Marras W, Ferguson S, Simon S. Three dimensional dynamic motor performance of the normal trunk. Int J Ind Ergonom 1990;6:211–24.
27. Marras W, Wongsam P. Flexibility and Velocity of the normal and impaired lumbar spine. Arch Phys Med Rehabil 1986;67:213–7.
28. Marras W, Fathallah F, Miller R, Davis S, Mirka G. Accuracy of a three-dimensional lumbar motion monitor for recording dynamic trunk motion characteristics. Int J Ind Ergonom 1992;9:75–87.
29. Marras W, Parnianpour M, Ferguson S, Kim J, Crowell R, Simon S. Vienna Physical Medicine Award 1993. Quantification and classification of low back disorders based on trunk motion. Eur J Phys Med 1993;3:218–35.
30. Marras W, Parnianpour M, Ferguson S, et al. The classification of anatomic- and symptom- based low back disorders using motion measure models. Spine 1995;20:2531–46.
31. Marras W, Parnianpour M, Kim J, Ferguson S, Crowell R, Simon S. The effect of task asymmetry, age and gender on dynamic trunk motion characteristics during repetitive trunk flexion and extension in a large normal population. IEEE Trans Rehabil Eng 1994;2:137–46.
32. Marras W, Lewis K, Ferguson S, Parnianpour M. Sincerity of effort during dynamic trunk motions. Spine (in press).
33. Mayer T, Kondraske G, Mooney V, Carmicheal T, Butsch R. Lumbar myoelectric spectral analysis for endurance assessment a comparison with deconditioned patients. Spine 1989;14:986–91.
34. Mayer T, Tencer A, Kristoferson S, Mooney V. Use of noninvasive techniques for quantification of spinal range-of-motion in normal subjects and chronic low back dysfunction patients. Spine 1984;9:588–95.
35. McClure P, Esola M, Schreier R, Siegler S. Kinematic Analysis of lumbar and hip motion while rising from a forward, flexed position in patients with and without a history of low back pain. Spine 1997;22:552–8.
36. McIntyre D, Glover L, Conino C, Seeds R, Levene J. A comparison of the characteristics of preferred low back motion of normal subjects and low-back-pain patients. J Spinal Disord 1991;4:90–5.
37. Mellin G, Hurri H, Harkapaa K, Jarvikoski A. A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Scand J Rehabil Med 1989;21:91–5.
38. Newton M, Thow M, Somerville D, Henderson I, Waddell G. Trunk strength testing with iso-machines. Part 2: Experimental evaluation of the cybex II back testing system in normal subjects and patients with chronic low back pain. Spine 1993;18:812–24.
39. Paquet N, Malouin F, Richards C. Hip–spine movement interaction and muscle activation patterns during sagittal trunk movements in low back pain patients. Spine 1994;19:596–603.
40. Peach J, McGill S. Classification of low back pain with the use of spectral electromyogram parameters. Spine 1998;23:1117–23.
41. Ripley B. Pattern Recognition and Neural Network. Cambridge: Cambridge University Press. 1996.
42. Robinson M, Cassisi J, O'Connor R. Lumbar image during isotonic exercises: Chronic low back pain patients vs controls. J Spinal Disord 1992;5:8–15.
43. Roy S, De Luca C, Emley M, Buijs R. Spectral electromyographic assessment of back muscles in patients with low back pain undergoing rehabilitation. Spine 1995;20:38–48.
44. Sherman R. Relationships between strength of low back muscle contraction and reported intensity of chronic low back pain. Am J Phys Med 1985;64:190–200.
45. Spitzer W, LeBlanc F, Dupruis M. Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. Spine 1987;12:S1–S59.
46. Spratt K, Lehman T, Weinstein J, et al. A new approach to low back examination: behavioral assessment of mechanical signs. Spine 1990;15:96–102.
47. Teasell R, Harth M. Functional restoration. Returning patients with chronic low back pain to work: Revolution or fad? Spine 1996;21:844–7.
48. Triano J, Schultz A. Correlation of objective measure of trunk motion and muscle function with low back disability ratings. Spine 1987;12:561–65.
49. U. S. Bureau of Disability Insurance. Disability Evaluation Under Social Security. A Handbook for Physicians. Washington DC: U. S. Government Printing Office, 1970.
50. Wadell G, Allan D, Newton M. Clinical evaluation of disability in low back pain. In: Frymoyer JW, ed. The Adult Spine: Principles and Practice. 2nd ed. Philadelphia: Lippincott–Raven, 1997:171–83.

Address reprint requests to

William S. Marras, PhD
Professor and Director
Biodynamics Laboratory
The Ohio State University
1971 Neil Avenue
Columbus, OH 43210