ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT

Manual Therapy ∎ (∎∎∎∎) ∎∎∎–∎∎∎

MANUAL THERAPY

www.elsevier.com/locate/math

Original article

Subjective and objective descriptors of clinical lumbar spine instability: A Delphi study

Chad Cook[a,*], Jean-Michel Brismée[b], Phillip S. Sizer Jr[b]

[a]Duke University Medical Center 3907, Durham, NC 27710, USA
[b]Department of Rehabilitation Sciences, Texas Tech University Health Sciences Center, Lubbock, TX 79430, USA

Received 12 May 2004; received in revised form 3 December 2004; accepted 4 January 2005

Abstract

Accurate ability to diagnose lumbar spine clinical instability is controversial for numerous reasons, including inaccuracy and limitations in capabilities of radiographic findings, poor reliability and validity of clinical special tests, and poor correlation between spinal motion and severity of symptoms. It has been suggested that common subjective and objective identifiers are specific to lumbar spine clinical instability. The purpose of this study was to determine if consensual, specific identifiers for subjective and objective lumbar spine clinical instability exist as determined by a Delphi survey instrument. One hundred and sixty eight physical therapists identified as Orthopaedic Clinical Specialists (OCS) or Fellows of the *American Academy of Orthopaedic Manual Physical Therapists* participated in three Delphi rounds designed to select specific identifiers for lumbar spine clinical instability. Round I consisted of open-ended questions designed to provide any relevant issues. Round II allowed the participants to rank the organized findings of Round I. Round III provided an opportunity to rescore the ranked variables after viewing other participant's results. The results suggest that those identifiers selected by the Delphi experts are synonymous with those represented in related spine instability literature and may be beneficial for use during clinical differential diagnosis.

Keywords: Clinical instability; Lumbar spine; Delphi; Physical therapy

1. Introduction

Since 1987, several low-back diagnostic classification systems have been created, each designed to categorize patients with low-back pain into homogenous subgroups for better clinical management decisions (McKenzie, 1981; Bernard and Kirkaldy-Willis, 1987; Delitto et al., 1995; Moffroid et al., 1994; Werneke and Hart, 2004). Data suggest that patients who are treated based on diagnosis or symptom-specific individual categorization into a classification system have superior outcomes than those who are not (Delitto et al., 1995; Deyo, 1993; Erhard et al., 1994; Riddle, 1998).

One controversial diagnostic classification is lumbar spine instability. This classification is controversial because the pathomechanical behaviour of lumbar spine instability is ambiguous and poorly defined (Paris, 1985; Dvorak et al., 1991; Panjabi, 1992; Lindgren et al., 1993; Cattrysse et al., 1997; Fritz et al., 1998; Olson and Joder, 2001). Regardless of clinical or radiographic (static or dynamic) test methods used, there is little evidence to relate the pathophysiological condition of spine instability with severity of verbal and objective symptoms (Farfan and Gracovetsky, 1984; Dupuis et al., 1985; Boden and Wiesel, 1990; Lindgren et al., 1993; Sihvonen et al., 1997). There may be several reasons for this finding. First, traditional radiographic measurement may suffer from errors during measurement of movements less than 5 mm (Shaffer et al., 1990; Harrison et al., 1998), which are frequently observed with spinal

*Corresponding author. Tel.: +1 915 335 5370; fax: +1 915 335 5365.
E-mail address: cookttu@yahoo.com (C. Cook).

1356-689X/$ - see front matter
doi:10.1016/j.math.2005.01.002

instability. Second, despite numerous attempts at standardization, quantification of the ''normal'' spinal motion is not yet precisely defined (Boden and Wiesel, 1990; Panjabi et al., 1994). Since asymptomatic spines show considerable variability of segmental motion amplitude, a diagnosis based on segmental motion alone may result in misleading or erroneous conclusions (Dupuis et al., 1985; Dvorak et al., 1991; Ogon et al., 1997). Third, traditional radiographic films concentrate on end range movements while instability is often present in midrange movements where one observes the neutral zone (Posner et al., 1982; Farfan and Gracovetsky, 1984; Dupuis et al., 1985; Frymoyer et al., 1990). Lastly, there is currently a lack of reliable and valid clinical assessment tools or special tests for spine instability (Nachemson, 1991; Taylor and O'Sullivan, 2000).

Spine instability may be best classified into two categories, (1) radiologic appreciable instability and (2) clinical instability. Radiologic appreciable instability reflects marked disruption of passive osseoligamentous anatomical constraints (Dupuis et al., 1985) and is typically diagnosed by appropriate radiographic measurements (Panjabi, 1992). Clinical instability is more challenging to diagnose and may involve discrepancies in radiographic findings (Panjabi, 1992). Theoretically, clinical dynamic stabilizers include the neural feedback systems, muscles and tendons of the spinal column and comprise force or motion transducers that include muscle spindles and Golgi tendon organs that exhibit proprioceptive and kinesthetic neural properties. Clinical instability commonly demonstrates subtle quantifiable clinical features (Niere and Torney, 2004) with negative or inconsistent findings during traditional radiographic analysis (Hayes et al., 1989; Takayanagi et al., 2001; Iguchi et al., 2003).

Despite indistinct findings for clinical instability, several authors have suggested that commonalities exist in subjective and objective descriptors (Paris, 1985; Lundberg and Gerdle, 2000; Taylor and O'Sullivan, 2000). Common subjective reports include ''giving way'', ''locking'', (O'Sullivan, 2000) or sensations of ''slipping out'' during the normal demands of spinal mobility (Ogon et al., 1997). Afflicted individuals may complain of a ''catch'' sensation during extension motions of the low back when returning from a flexed posture (Taylor and O'Sullivan, 2000) or the necessity to ''twist the back into position'' (Paris, 1985). Selected authors report instability is associated with pain immediately upon sitting, and relieved through standing or temporary movement (Paris, 1985; Maigne et al., 2003). Others suggest the presence of ''through range pain'' and a painful arc during active motions (Kirkaldy-Willis and Farfan, 1982; O'Sullivan, 2000). Individuals with clinical instability may report single or multiple causal incidents and frequently report poor outcomes with general resistance exercise programs, spinal manipulation, and mobilization-based treatments (Kirkaldy-Willis and Farfan, 1982; O'Sullivan, 2000).

During the physical examination, specific patterns are often associated with clinical instability. Investigators have reported the patient's inability to stand erect without the assistance of the hands (Kirkaldy-Willis and Farfan, 1982; O'Sullivan, 2000). Patients often show signs of unexpected movements, accelerations and small jerks that occur intersegmentally, and hesitations or giving way during active motion (Ogon et al., 1997). Another common clinical manifestation is increased global muscular tone (Dvorak et al., 1991; Panjabi, 1992), representing the body's attempt to stabilize the hypermobile segment.

1.1. Purpose of the study

The purpose of this study was to establish consensus regarding the common subjective and objective symptoms associated with clinical instability of the spine. Clinical instability is a challenging yet common disorder seen by physical therapists, which lacks a definitive method for diagnosis (Taylor and O'Sullivan, 2000). By using a Delphi survey instrument, expert practitioners defined common identifiers of lumbar clinical instability. In turn, the consensus agreement could be valuable for enhancing the clinician's differential diagnosis and appropriate classification of spinal instability.

2. Materials and methods

2.1. Study design

This investigation implemented a Delphi survey instrument that incorporated both a respondent group and a work group. The respondent group was comprised of the target population of experts used within this study. The work group was comprised of those investigators who summarized the returned data from Round I and redesigned the follow-up instruments.

2.2. Respondent group

The population pool selected for the study consisted of volunteers from two ''expert'' categories. First, all Board Certified Orthopaedic Clinical Specialists (OCS) from the *American Physical Therapy Association* (APTA) who identified cervical and lumbar dysfunction as their primary practice specialty, were targeted as population pool number one. The APTA proposes that designation of orthopaedic board certification through the APTA depicts a clinician with ''knowledge, skill and experience exceeding that of the physical therapist at entry to the profession and unique to the specialized area of practice'' (APTA, 2004). Second, *Fellows of the American Academy of Orthopaedic Manual Physical Therapists* were additionally targeted, based on their clinical expertise obtained through residency. For completion of a fellowship, the AAOMPT requires a physical therapist to complete a credential fellowship program in orthopaedic manual physical therapy or demonstrate the equivalent of competence by successfully passing a portfolio review process and oral/practical examination (AAOMPT, 2004). Since 2001, the APTA and AAOMPT have merged their credentialing processes for post-professional orthopaedic residency and manual therapy fellowship training to establish clear standards for professional development and credentialing. All targeted participants were contacted using traditional mail-outs and e-mails (when possible), and were pooled into one group upon agreement to participate.

2.3. Work group

This group was comprised of three individuals, including the primary investigator and two individuals experienced in qualitative research. All three work group members were Physical Therapists with at least 14 years of both research and clinical experience in orthopaedic manual therapy.

2.4. Instrumentation

For the present study, a Web-based three-round Delphi survey instrument was generated. A typical Delphi survey instrument consists of three rounds of questionnaires that respondents consecutively answer in a timely fashion (Binkley et al., 1993). Generally, at the end of the third round, consensus is generated among respondents. Round I consisted of a series of open-ended questions designed to identify specific issues relevant to the overall survey topic (Cleary, 2001). Round I of the instrument included questions regarding basic demographic information and open-ended questions inquiring about subjective identifiers and objective clinical findings in patients with lumbar instability. After defining lumbar instability, the first open-ended question in Round I asked respondents to identify the subjective factors they deemed were associated with lumbar spine instability. The second open-ended question asked respondents to identify objective or physical factors they deemed were associated with lumbar spine instability.

Round II of the instrument was comprised of a list of descriptor statements that defined each subjective and objective factor that were constructed from the work group's qualitative analysis of Round I. Respondents were asked to use Likert-type grading scales to score each of these factors in terms of whether each is related to lumbar spine instability.

Round III of the instrument was comprised of the same list and grading scale as Round II, with additional tables and graphs for each descriptor statement demonstrating the descriptive statistical score outcomes for each factor. The graphic information identified the percentage of total respondents that selected each possible score for the given item in Round II. The respondents were instructed to rescore each item with the scale after viewing the scoring results from Round II.

2.5. Procedure

Expedited approval of the experiment was granted by the Institutional review Board at *Texas Tech University Health Sciences Center*. Invitations to Round I of the instrument were automatically distributed through E-mail to all potential respondents and provided a Web link to the Web-based consent form. Respondents that did not answer the request for participation were e-mailed a reminder notice to encourage compliance using a method suggested by Dillman (2000). Two consecutive follow-up reminders were delivered at 10 and 20 days after the initial questionnaire was provided, respectively (Lopopolo, 1999; Pesik et al., 1999; Vaughan-Williams et al., 1999).

After respondents completed Round I, the *WebSurveyor* program (WebSurveyor, 2004) automatically downloaded response data onto a spreadsheet for work group analyses. Work group members coded each single subjective and objective factor data entry on different levels in accordance with suggestions provided by Berg (2001). First, members performed a quantitative analysis whereby data entries were coded based on similar words or phrases (known as ''literal coding''), where word groupings were used to create categories that represented the group content with names such as ''changes with manipulation.'' Work group members used a thesaurus when words demonstrated similarity but were unfamiliar to the member. Once these categories were established, then the remaining data entries were coded in a qualitative fashion, whereby data were entered into the specific categories based on similar meanings and contexts. When a data entry did not fit into previously created categories, then a new category was formed.

The coding process was first conducted for the data entries received under the *Subjective Factors* heading, and then for the data under the *Objective Factors* heading. The work group moved a data entry when it appeared more appropriate under the other heading (i.e. subjective vs. objective). Upon completion of the individual coding process, the work group convened to begin the group coding process, whereby each individual data entry was discussed and entered into a collective category. However, each data entry was coded into a collective category only when all three work group members unanimously (100%) agreed on the category assignment. If the members persisted in disagreement, then the entry was discarded.

After completing the group coding process, the work group created descriptor statements that summarized the content within each collective category. For example the work group used the following title "Frequent clicking, grinding, crepitation, and popping during movement" to represent a descriptor category that included data entries describing any of those behaviours reported by a patient. The descriptors were used to compose Round II of the Delphi instrument.

The purpose of this Round II was to allow informants to: (1) review the categories of responses from Round I for the clarification and correction of terminology; and (2) identify categories of responses they considered the most important with respect to the research (Lopopolo, 1999). Invitations to Round II of the instrument were automatically distributed through e-mail to all respondents from Round I, providing the respondents with a Web link to Round II. After viewing each descriptor under each heading in Round II, respondents were asked to score the importance of the descriptor using the following scale:

1 = *Strongly Agree*; the selected identifier has a very strong relationship with lumbar spine instability
2 = *Agree*; the selected identifier is related to lumbar spine instability
3 = *Undecided*; uncertainty of the relationship of the selected identifier and lumbar spine instability
4 = *Disagree*; the selected identifier is not related to lumbar spine instability
5 = *Strongly Disagree*; there is absolutely no relationship whatsoever with the selected identifier and lumbar spine instability

After respondents completed Round II, the *WebSurveyor* program was again used to download response data automatically. Tallies of respondents' scoring were then graphically represented based on the percentage of respondents who selected each score from the previously described scales.

Invitations to Round III of the instrument were automatically distributed in a similar fashion to Round II, once again providing a Web link to Round III. Respondents were asked to rescore each descriptor after viewing the descriptor statement along with its corresponding graph. Respondents used the previously described scale to rescore each descriptor after viewing the scoring distribution produced for that descriptor in Round II. After respondents completed Round III, the *WebSurveyor* program automatically downloaded response data as before, where data under each descriptor statement were sorted and tallied. Tallies of respondents' scoring were then used for statistical analyses.

2.6. Data analysis

For each descriptive identifier, the scores were divided from Round III into two categories: The tally of scores "*Strongly Disagree*" and "*Disagree*" represented total percentage of scores in the "Not Related" category, meaning that the subjective or objective factor is not important for the diagnosis of instability. Conversely, the tally of scores in the "*Strongly Agree*" and "*Agree*" categories were placed in the "Related" category, meaning that the subjective or objective factor is important for that diagnosis. Consensus was established if 75% or greater of the respondents (Binkley et al., 1993) scored the subjective or objective factor as "Not Related" (CNR) or "Related" (CR). Fig. 1 provides an example of a consensus-scoring tally.

If the tally for "Not Related" or "Related" was between 50% and 74.9%, then consensus was not established and a decision was forced between Near-consensus (Triezenberg, 1997) and Undecided. To arrive at the answer to the decision between Near-consensus and Undecided, a logic analysis was conducted. If the tally for "*Strongly Agree*" and "*Agree*" was greater than the tally for "*Agree*" and "*Disagree*", then the descriptor was labelled as "Near-consensus, Related (NCR)". Similarly, if the tally for "*Strongly Disagree*" and "*Disagree*" was greater than the tally for "*Agree*" and "*Disagree*", then the descriptor was labelled as "Near-consensus, Not Related". However, if the tally for "*Agree*" and "*Disagree*" was greater than the tally for "*Strongly Agree*" and "*Agree*" or the tally for

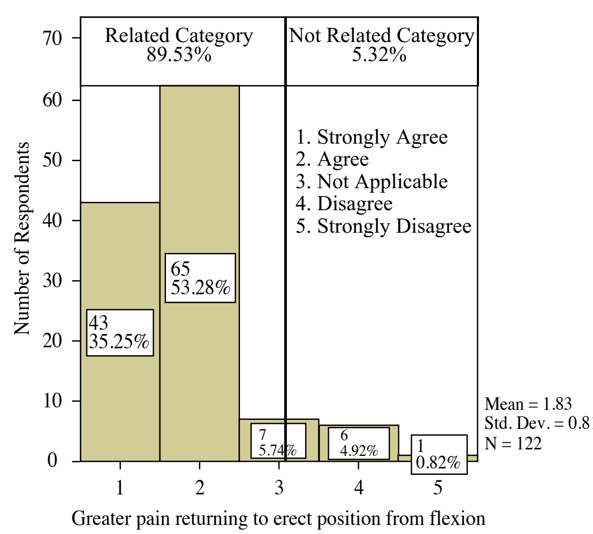

Fig. 1. Example of a consensus-scoring tally indicating consensus or not. The text represents the numeric value associated with the Likert-like value. Because over 75% of scores were in Likert scores 1 and 2, it was deemed that consensus was reached.

"*Strongly Disagree*" and "*Disagree*," then the descriptor was labelled as "Undecided (U)".

After consensus was established, the subjective and objective factors were ranked. This was accomplished by first assigning a composite score to each factor. The composite score was determined using the following formula:

$$\text{Composite Score} = (n1 \times 5) + (n2 \times 4) + (n3 \times 3) + (n4 \times 2) + (n5 \times 1),$$

where the subjective or objective factors were tallied as: $n1$ is the number of respondents who scored the factor as "*Strongly Agree*", $n2$ the number of respondents who scored the factor as "*Agree*", $n3$ the number of respondents who scored the factor as "*Undecided*", $n4$ the number of respondents who scored the factor as "*Disagree*", $n5$ the number of respondents who scored the factor as "*Strongly Disagree*".

A graphic example of this composite score tally is presented in Fig. 2. Here, the composite score value was derived from the tally of scores for each descriptor statement. Consequently, the skill in Fig. 2 was assigned a composite score of 509. The composite scores were subsequently used for ranking under each heading (subjective or objective factors) with the highest score representing the most important factor under each heading.

In a Delphi design, the respondents rank composite scores both without (Round II) and with (Round III) graphic feedback from the other respondents, thus it is expected that some changes will occur between rounds. Using *Megastat* version 9.0, a Mann–Whitney U ($\alpha = 0.05$) was used to determine if a meaningful difference between ranked scores between Rounds II and III is present for both subjective and objective factors.

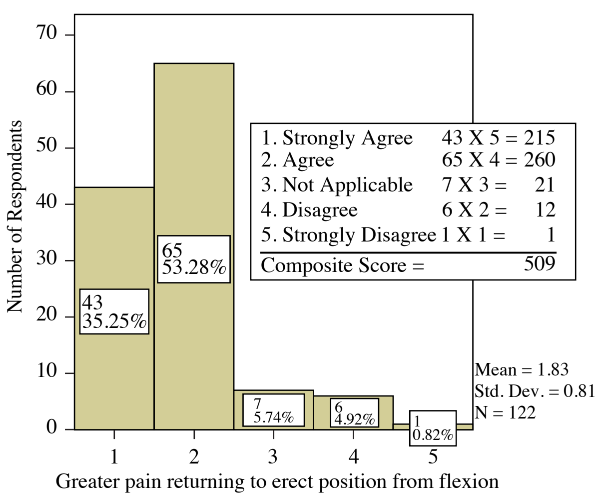

Fig. 2. Composite score tally sheet. The text bar represents the calculations associated with composite score ranking. The total composite score is then compared to other descriptors.

3. Results

3.1. Results from Round I

One thousand one hundred and eleven orthopaedic certified specialists from the APTA and 334 Fellows in the AAOMPT were solicited for participation in the present study. Of the 1111 individuals, 92 were not accessed due to incorrect electronic mail address, server difficulties, or relocation without a new address. One hundred and sixty-eight individuals (11.6% return rate) chose to respond to Round I. Participant demographic and clinical characteristics are outlined in Table 1.

The work group agreed to eliminate data points when: (1) The members could not reach 100% consensus regarding the appropriate descriptor category to which a data point belonged; (2) a data point could fit into more than one descriptor category without any persuasion towards any particular category; and/or (3) a data point appeared irrelevant, incomprehensible, or incomplete. Group coding produced 33 descriptors under the *Subjective Factors* heading, while 28 descriptors were produced under the *Objective Factors* heading (see Appendix A and B).

Table 1
Respondent demographics, credentials, work setting, and report of the first, second and third most influential manual therapy models that have influenced their personal clinical practice experiences

Age	Mean = 43, range 27–61 years Missing values = 3	
Gender	Male = 96 Female = 72 Missing values = 4	
Credentials	FAAOMPT = 66 OCS = 78 Both = 49 Missing values = 28	
Years of experience	Mean = 17.5 years, range 3–39 years Missing values = 3	
Work setting	107 > 50% of clinical time in non-hospital-based outpatient 38 > 50% of clinical time in hospital- based outpatient Missing values = 27	
Reported background	Grimsby	4.12%
	Kaltenborn	8.24%
	Maitland	24.12%
	McKenzie	14.71%
	NA	0%
	NAIOMPT[a]	7.65%
	Osteopathic	19.41%
	Other	8.24%
	Paris	12.35%
	Winkel	1.18%

[a]North American institute of orthopaedic manual therapy.

3.2. Results from Rounds II and III

Twenty-eight subjects failed to leave e-mail contact information during Round I, thus, only 140 of the 168 respondents were contacted for participation in Round II. One hundred and thirty-three of the 140 respondents (95% retention rate between Rounds I and II; 9.2% overall response rate) completed Round II, and 122 (72.6%) completed Round III producing a 92% retention rate between Rounds II and III and an overall response rate of 8.4%. The total composite score tallies for Rounds II and III are reported in Appendix A for *Subjective Factors* and Appendix B for *Objective Factors*. Maximum and minimal composite score values for Round III were 610 for maximum agreement and 122 for maximum disagreement. For Round III, 15 *Subjective Factors* reached consensus as ''Related'' (CR) with lumbar spine instability, while two reached ''NCR''. Additionally, one subjective factor reached ''Consensus Not Related'' (CNR) and 15 were deemed undecided (U). For the *Objective Factors*, 14 reached consensus as being ''Related'' with lumbar spine instability, while two reached ''Near Consensus Related'', three reached ''CNR'' and nine were deemed as ''Undecided''.

Factor rank outcomes are reported in Appendix A and B. *Reports feelings of ''giving way'' or back ''giving out''* ranked as the subjective factor that is most related with lumbar spine clinical instability. *Self-Manipulator who feels the need to frequently crack or pop the back* ranked second, followed by *Frequent bouts or episodes of symptoms. Pain through the range of motion* (i.e. *through range pain*), *Intolerance of prone position*, and *Spine instability does not exist* ranked as the three subjective factors that are least related to lumbar spine instability.

Overall, Poor lumbopelvic control, including segmental hinging or pivoting with movement, as well as Poor proprioceptive function ranked as the objective factors that were most related to lumbar spine instability, followed by Poor coordination/neuromuscular control, including juddering or shaking. The third most related objective factor was Decreased strength and endurance of local muscles at the level of segmental instability. Additionally, the three objective factors that were determined to be least related with lumbar spine instability included Non-objectifiable: Segmental instability cannot be objectified in the clinic; Unresponsiveness to treatment, including manual techniques and exercise; and Segmental instability does not exist.

Finally, no significant difference in composite score rankings was detected through data analysis for Rounds II and III in the *Subjective Factors* ($P = 0.43$, $df = 33$, $U = 482$) or *Objective Factors* ($P = 0.36$, $df = 28$, $U = 336$). This indicates that the influence of seeing the other respondent selections did not significantly alter the ranks in Round III.

4. Discussion

The Delphi technique is a series of sequential questionnaires designed to distill and obtain the most reliable consensus from a group of experts (Powell, 2003). The method is useful in situations where frequent clinical or practical judgments are encountered, yet incomplete empirical evidence exists to provide evidence-based decision-making (Fink et al., 1984; Dawson and Barker, 1995; Powell, 2003). At present, clinical detection of lumbar spine instability using pathoanatomic, radiologic, and clinical assessment has inherent limitations (Dupuis et al., 1985; Dvorak et al., 1991; Nachemson, 1991; Lindgren et al., 1993; Ogon et al., 1997).

Adler and Ziglio (1996) state that in the absence of complete information the health care provider may wait until they have enough information to create an adequate theory, or they may make the most of the available information and use this knowledge for the best possible consequence. Studies of clinical reasoning identify that expert clinicians recognize inconsistencies during an examination and have the capacity to combine clusters of information together into workable sets, based on past-experience and cooperative decision-making (Jensen et al., 2000). The investigations suggest that judicious use of the Delphi findings may contribute to a growing pool of data for identification of clinical instability. Thus, by using the clusters of identifiers proposed within the Delphi consensus, practitioners may glean additional information for successful assessment of clinical spine instability.

4.1. Subjective findings

Recently, a questionnaire completed by patients diagnosed with lumbar instability described back pain symptoms as ''*recurrent*'' (70% of participants), ''*constant*'' (55% of participants), ''*catching*'' (45% of participants), ''*locking*'' (20% of participants), ''*giving way*'' (20% of participants), and/or ''*accompanied by a feeling of instability*'' (35% of participants) (Taylor and O'Sullivan, 2000). The Delphi participants in the present study reported consensus-related factors of *giving way* and *giving out*, *frequent bouts*, and a *condition that is progressively worsening*. They also identified the frequent need to *self manipulate* their spine as a pain control mechanism. The Delphi participants also recognized common subjective complaints during postures, movements, or activities.

The Delphi participants outlined history of painful locking or catching during twisting or bending of the spine, pain during transitional activities, pain during return from a flexed position, pain during sudden or trivial activities, difficulty with unsupported sitting, and pain that is worsened with sustained postures as signs of clinical lumbar instability. O'Sullivan (2000) identified pain during "forward bending" (75% of participants), "sudden unexpected movements" (75% of participants), "returning to an upright position from forward bending" (65% of participants), "lifting" (65% of participants), and "sneezing" (60% of participants) as signs of clinical lumbar instability. These findings are very similar to those agreed upon by the expert panel used in the present study and suggest that historical information gathered by the clinician in patients with lumbar spinal symptoms provides essential data for the appropriate differential diagnosis of clinical lumbar instability.

4.2. Objective findings

Identification of muscle dysfunction, motor control abnormalities, and strength losses were the strongest identifiers of clinical lumbar instability selected by the Delphi participants. The descriptors *poor lumbopelvic control including segmental hinging or pivoting with movement*, *muscle guarding/spasm*, *poor coordination/neuromuscular control including juddering or shaking*, and *decreased strength and endurance of local muscles at level of segmental instability* were the top three component scores. Investigators have suggested that clinicians may be capable of discriminating these criteria in an accurate fashion (Panjabi, 1992; Hodges and Richardson, 1996), which was verified in recent studies (Hides et al., 1996; O'Sullivan, 2000; Radebold et al., 2000; van Dieen et al., 2003). Such discrimination could be clinically useful, as a clinician's analysis of specific muscular anatomical action can influence one's clinical assessment. Additionally, appropriate identification of motor impairment allows targeting of specific muscles groups for active recovery (O'Sullivan, 2000).

The Delphi participants identified *segmental mobility including pain provocation techniques* as specific identifiers to clinical spine instability. Additionally, this group recognized *excessive motion of one of two segments during flexion extension*, *hypermobility during PA spring testing*, *positive pain during a PA spring test*, and *hypomobile adjacent segments* as consensus identifiers. This finding that manual palpation is effective in detection of instability, simply by determining if motion is greater than that found with hypermobility has been suggested by others (Kirkaldy-Willis and Farfan, 1982; Paris, 1985). Little evidence exists to support the reliability of palpation mechanisms for spinal instability assessment (Gonnella et al., 1982; Dupuis et al., 1985; Olson et al., 1998; Panjabi et al., 1998). The sensitivity, specificity, and predictive value of the numerous clinical examination techniques, including special tests, have not been fully recognized (Nachemson, 1991; Cattrysse et al., 1997). Most manual instability assessment methods are finite, require very skilled assessment and have not been substantiated by simultaneous diagnostic measurement (Maigne et al., 2003). Some spinal instability conditions are not easily quantified in the absence of an externally loaded position such as standing (Dupuis et al., 1985), and the majority of instability tests are performed in an unloaded position such as supine, side-lying or prone. Additionally, many of the previously studied segmental tests were performed in isolation and may improve when combined with other clusters of information.

The Delphi participants also recognized various observed motion disparities during position changes and pain during certain postures. These identifiers included *pain with sustained postures and positions*, *Gower's sign*, and *poor posture and postural deviations including lateral shift*, all similar to findings found by other authors (Kirkaldy-Willis and Farfan, 1982; O'Sullivan, 2000). Previous authors have indicated that acute instability cases often exhibit retroscoliosis (Kirkaldy-Willis and Farfan, 1982; Boden and Frymoyer, 1997). Moreover, Maigne et al. (2003) indicate that pain is often present upon immediate sitting, relieved once the individual returns to standing. O'Sullivan (2000) reported that sustained sitting, standing, and sustained slight flexion in standing were the most commonly identified postural complaints. Others have suggested the occurrence of intermittent neurological symptoms such as depressed reflexes and positive dural signs (Boden and Frymoyer, 1997) findings that were not identified by the Delphi participants.

Selected authors have suggested that aberrant spinal motions during physiological movements that produce catching and disruptions of a normal smooth arc of motion are indicative of spine instability (Kirkaldy-Willis and Farfan, 1982; Nachemson, 1985). Using dynamic motion analysis, Ogon et al. (1997) quantified hesitation, giving way, and a "jerk" during active motion, which were clinical observations that were previously unverified. Patients with spine instability often demonstrate signs of unexpected segmental movements, accelerations and decelerations (Ogon et al., 1997). The Delphi participants recognized *aberrant movement including lateral shift changes* and *motion disparity during weight bearing and non-weight bearing and during active range of motion and passive range of motion* as related to clinical instability in a fashion similar to previously published findings.

Selected few respondents suggested that lumbar spine instability does not exist, although the literature does strongly suggest the existence of such a condition, albeit in a controversial fashion. Conversely, the majority of respondents agreed that the condition does exist and that many of the identifiers serve as clinical assessment tools commonly used in daily practice for the recognition of clinical instability. This consensus was consistent with the previous reports in the literature. Two plausible explanations exist for this agreement. First, the Delphi participants may be well informed of the literature and of specific objective identifiers of clinical lumbar instability reported in previous studies. Second, the results reported in the studies that investigated lumbar spinal instability could be strongly associated with what expert clinicians have empirically observed during day-to-day practice.

4.3. Limitations

The Delphi instrument is a qualitative analysis and is immune from the sampling requirements of a randomized design. Fewer than 12% of the targeted population responded to initial recruitment. There may be several reasons for the low response rate. First, e-mail annual response rates for surveys dropped consistently from 1992 to 2000 (Sheehan, 2001). Second, it is estimated that the average e-mail user receives 39 unsolicited e-mails each day and has prompted many users to create multiple e-mail addresses, including "bulk addresses" unsolicited mail (Sheehan, 2001). Third, this study used an e-mail system that does not report when e-mails are no longer in service, thus chance exists that the introductory e-mails did not arrive at all the potential 1015 eligible OCS participants. Although a Delphi instrument is appropriate for sample sizes as small as 10–12 participants, it is essential that the experts selected are truly representative of the most talented clinicians in this targeted field. The assumption of this study is that the criteria required for OCS and fellowship status within the AAOMPT do meet those specifications.

5. Conclusion

This Delphi investigation was designed to identify common subjective and objective identifiers for lumbar spine instability. The findings suggest that those identifiers selected by the Delphi experts are similar to those reported in spine-related instability literature, suggesting that those identifiers are specifically associated with lumbar spine clinical instability and may be beneficial for clinical differential diagnosis. Future research should examine patterns or clusters of identifier findings in patients *suspected* of suffering from clinical instability. The Delphi information could be used as a scale to determine a threshold point for analysis. Additionally, a factor analysis could identify themes of instability suggested by the Delphi group. Lastly, spine instability special tests require validation with populations with measurable instability, specifically since clinical spine instability most likely exhibits more difficulty in definitive assessment.

Acknowledgements

This study was approved by the *Texas Tech Health Sciences Center* Institutional Ethics Review Board and financially supported by the 2003 Steens/USA Grant.

Appendix A

Subjective factors of consensus and rank outcomes for clinical lumbar spine instability, listed in descending rank.

Descriptor	Round III consensus status	Round II composite scores	Round III composite scores
Reports feelings of "giving way" or back "giving out"	CR	501	527
Self manipulator who feels the need to frequently crack or pop the back	CR	483	524
Frequent bouts or episodes of symptoms	CR	518	523
History of painful catching or locking during twisting or bending of the spine	CR	496	521
Pain during transitional activities (e.g. sit to stand)	CR	484	510
Greater pain returning to erect position from flexion	CR	493	509
Pain increased with sudden, trivial, or mild movements	CR	496	504
Difficulty with unsupported sitting and better with supported backrest	CR	477	500
Worse with sustained postures and a decreased likelihood of reported static position that is not painful	CR	470	495
Condition is progressively worsening (e.g. shorter intervals between bouts)	CR	471	490
Long-term, chronic history of disorder	CR	457	478
Temporary relief with back brace or corset	CR	463	478
Reports frequent episodes of muscle spasms	CR	482	474

Frequent clicking, grinding, crepitation, and popping during movements	NCR	453	466
Dramatic but temporary relief with manipulation	NCR	436	465
Fear and decreased willingness to move	CR	461	464
Reports of previous back injury or trauma	CR	442	457
Record of poor improvement with past treatment interventions	U	446	454
Reports sleep disturbances including frequent position changes during sleep	U	440	444
Inconsistent, non-specific symptoms such as pain which alternates from side to side	U	422	440
Frequently feel "tight" or "stiff" and needs to stretch a lot	U	439	437
Inconsistency of symptoms (i.e. pain is not provoked upon command)	U	425	433
History of predisposing sports or labour (i.e. gymnastic, weight lifting)	U	434	432
Transient neurological symptoms	U	425	424
Pain with weight bearing activities. Relief with non-weight bearing activities	U	403	412
Pain which is provoked by ADLs	U	409	409
Pain worse at end or extreme ranges	U	400	396
Pain that is better in the A.M. and worse as the day progresses	U	403	395
Reports of "central" or "centralized" symptoms in low-back pain	U	400	382
History of pregnancy or oral contraceptive use	U	403	370
Pain through the range of motion (i.e. through range pain)	U	369	338
Intolerance of prone position	U	350	331
Spine instability does not exist	CNR	196	145

Definitions: CR = consensus related; NCR = near consensus related; CNR = consensus not related; U = undecided.

Appendix B

Objective factors of consensus and rank outcomes for clinical lumbar instability, listed in descending rank.

Descriptor	Round III consensus status	Round II composite scores	Round III composite scores
Poor lumbopelvic control, including segmental hinging or pivoting with movement, as well as poor proprioceptive function	CR	517	539
Poor coordination/neuromuscular control, including juddering or shaking	CR	488	537
Decreased strength and endurance of local muscles at level of segmental instability	CR	522	533
Aberrant movement, including changing lateral shift during AROM	CR	486	510
Pain with sustained positions and postures	CR	479	507
Gower's sign: Patient walks up thighs when returning from flexion	CR	492	503
Excessive motion of one of two segments during flexion-extension	CR	487	503
Decreased willingness or apprehension of movement	CR	491	494
Hypermobility during posterior–anterior (PA) Spring test	CR	473	493
Increased muscle guarding/spasm	CR	475	474
Poor posture and postural deviations that include lateral shift and changes in lordosis	CR	449	471
Positive spring test (PA provocation test)	CR	447	466
Frequent catching, clicking, clunking and popping heard during movement	CR	447	461
Motion disparity between weight bearing and non-weight bearing	NCR	442	460
Hypomobile adjacent segments	CR	457	460
Motion disparity between AROM vs. PROM	NCR	425	456
Pain with palpation, including interspinous space and ligament	U	428	446
Hypertrophic erector spinae	U	438	443
Palpable segmental position change	U	417	434

Prone instability test that includes passive segmental rotation	U	414	418
Positive radiographic evidence, including traction spurs	U	414	423
Inconsistent examination findings	U	393	396
Excessive active physiological ROM	U	406	396
Findings of overall, generalized laxity	U	372	382
Quadrant test painful bilaterally	U	340	312
Non-objectifiable: segmental instability cannot be objectified in the clinic	CNR	305	259
Unresponsiveness to treatment, including manual techniques and exercise	CNR	268	254
Segmental instability does not exist	CNR	310	152

Definitions: CR = consensus related; NCR = near consensus related; CNR = consensus not related. U = undecided.

References

Adler M, Ziglio E. Gazing into the oracle. Bristol, PA: Jessica Kingsley Publishers; 1996.

Berg B, editor. Qualitative research methods for the social sciences. 4th ed. Boston: Allyn and Bacon; 2001.

Bernard T, Kirkaldy-Willis W. Recognizing specific characteristics of nonspecific low back pain. Clinical Orthopedics 1987;217: 266–80.

Binkley J, Finch E, Hall J, Black T, Gowland C. Diagnostic classification of patients with low back pain: report on a survey of physical therapy experts. Physical Therapy 1993;73:128–55.

Boden S, Frymoyer J. Segmental instability. The adult spine: principles and practice. 2nd ed. Philadelphia: Lipponcott-Raven Publishers; 1997.

Boden S, Wiesel S. Lumbosacral segmental motion in normal individuals. Have we been measuring instability properly? Spine 1990;15:571–6.

Cattrysse E, Swinkels R, Oostendorp R, Duquet W. Upper cervical instability: are clinical tests reliable? Manual Therapy 1997; 2:91–7.

Cleary K. Using the Delphi process to reach consensus. Journal of Cardiopulmonary Physical Therapy 2001;1:20–3.

Dawson S, Barker J. Hospice and palliative care: a Delphi survey of occupational therapists roles and training needs. Australian Occupational Therapy Journal 1995;42:119–27.

Delitto A, Erhard R, Bowling R. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. Physical Therapy 1995;75:470–89.

Deyo R. Practice variations, treatment fads, rising disability. Do we need a new clinical research paradigm? Spine 1993;18:2153–62.

Dillman D. Mail and Internet surveys. The tailored design method. New York: Wiley; 2000 (p. 363).

Dupuis P, Yong-Hing K, Cassidy J, Kirkaldy-Wills W. Radiologic diagnosis of degenerative lumbar spinal instability. Spine 1985;10:262–76.

Dvorak J, Panjabi M, Novotny J, Chang D, Grob D. Clinical validation of functional flexion-extension roentgenograms of the lumbar spine. Spine 1991;16:943–50.

Erhard R, Delitto A, Cibulka M. Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome. Physical Therapy 1994;74:1093–100.

Farfan H, Gracovetsky S. The nature of instability. Spine 1984; 9:714–9.

Fink A, Kosecoff J, Chassin M, Brook R. Consensus methods: characteristics and guidelines for use. American Journal of Public Health 1984;74:979–83.

Fritz J, Erhard R, Hagen B. Segmental instability of the lumbar spine. Physical Therapy 1998;78:889–96.

Frymoyer J, Pope M, Wilder D. Segmental instability. In: Weinstein J, Wiesel S, editors. The lumbar spine. Philadelphia: WB Saunders; 1990. p. 612–36.

Gonnella C, Paris S, Kutner M. Reliability in evaluating passive intervertebral motion. Physical Therapy 1982;62:436–44.

Harrison D, Harrison D, Troyanovich S. Three-dimensional spinal coupling mechanics: part one. Journal of Manipulative Physiological Therapeutics 1998;21:101–13.

Hayes M, Howard T, Gruel C, Kopta J. Roentgenographic evaluation of lumbar spine flexion-extension in asymptomatic individuals. Spine 1989;14:327–30.

Hides J, Richardson C, Jull G. Multifidus muscle recovery is not automatic following resolution of acute first episode low back pain. Spine 1996;21:2763–9.

Hodges P, Richardson C. Inefficient muscular stabilization of the lumbar spine associated with low back pain. Spine 1996;21: 2640–50.

Iguchi T, Kanemura A, Kasahara K, Kurihara A, Doita M, Yoshiya S. Age distribution of three radiologic factors for lumbar instability: probable aging process of the instability with disc degeneration. Spine 2003;28:2628–33.

Jensen G, Gwyer J Shepard K, Hack L. Expert practice in physical therapy. Physical Therapy 2000;80:28–43.

Kirkaldy-Willis W, Farfan H. Instability of the lumbar spine. Clinical Orthopedics 1982;165:289–95.

Lindgren K, Sihvonen T, Leino E, Pitkanen M, Manninen H. Exercise therapy effects on functional radiographic findings and segmental electromyographic activity in lumbar spine instability. Archives of Physical Medicine and Rehabilitation 1993;74:933–9.

Lopopolo R. Hospital restructuring and the changing nature of the physical therapist's role. Physical Therapy 1999;79:171–85.

Lundberg G, Gerdle B. Correlations between joint and spinal mobility, spinal sagittal configuration, segmental mobility, segmental pain, symptoms and disabilities in female homecare personnel. Scandinavian Journal of Rehabilitative Medicine 2000; 32:124–33.

Maigne J, Lapeyre E, Morvan G, Chatellier G. Pain immediately upon sitting down and relieved by standing up is often associated with radiologic lumbar instability or marked anterior loss of disc space. Spine 2003;28:1327–34.

McKenzie R. The lumbar spine: mechanical diagnosis and therapy. Waikanae, New Zealand: Spinal Publications; 1981.

Moffroid M, Haugh L, Henry S, Short B. Distinguishable groups of musculoskeletal low back pain patients and asymptomatic control subjects based on physical measures of the NIOSH low back Atlas. Spine 1994;19:1350–8.

Nachemson A. Lumbar spine instability. A critical update and symposium summary. Spine 1985;10:290–1.

Nachemson A. Instability of the spine. Neurosurgery clinics of North America 1991;2:785–90.

Niere K, Torney S. Clinicians' perceptions of minor cervical instability. Manual Therapy 2004;9:144–50.
Ogon M, Bender B, Hooper D, Spratt K, Goel V, Wilder D, Pope M. A dynamic approach to spinal instability. Part I: sensitization of intersegmental motion profiles to motion direction and load condition by instability. Spine 1997;23:2841–58.
Olson K, Joder D. Diagnosis and treatment of cervical spine clinical instability. Journal of Orthopedic Sports Physical Therapy 2001;31:194–206.
Olson K, Paris S, Spohr C, Gorniak G. Radiographic assessment and reliability study of the craniovertebral side-bending test. Journal of Manual and Manipulative Therapy 1998;6:87–96.
O'Sullivan P. Lumbar segmental ''Instability'': clinical presentation and specific stabilizing exercise management. Manual Therapy 2000;5:2–12.
Panjabi M. The stabilizing system of the spine. Part 1 and Part 2. Journal of Spinal Disorders 1992;5:383–97.
Panjabi M, Oxland T, Yamamoto I, Crisco J. Mechanical behavior of the human lumbar and lumbosacral spine as shown by three-dimensional load-displacement curves. Journal of Bone and Joint Surgery, American 1994;76:413–24.
Panjabi M, Nibu K, Cholewicki J. Whiplash injuries and the potential for mechanical instability. European Spine Journal 1998;7:484–92.
Paris S. Physical signs of instability. Spine 1985;10:277–9.
Pesik N, Keim M, Sampson T. Do US emergency medicine residency programs provide adequate training for bioterrorism? Annals of Emergency Medicine 1999;34:173–6.
Posner I, White A, Edwards W, Hayes W. A biomechanical analysis of the clinical stability of the lumbar & lumbosacral spine. Spine 1982;7:374–89.
Powell C. The Delphi technique: myths and realities. Journal of Advanced Nursing 2003;41:376–82.
Radebold A, Cholewicki J, Panjabi M, Patel T. Muscle response pattern to sudden trunk loading in healthy individuals and in patients with chronic low back pain. Spine 2000;25:947–54.
Riddle D. Classification and low back pain: a review of the literature and critical analysis of selected systems. Physical Therapy 1998;78:708–37.
Shaffer W, Spratt K, Weinstein J. The consistency and accuracy of roentgenograms for measuring sagittal translation in the lumbar vertebral motion segment: an experimental model. Spine 1990;15:741–50.
Sheehan K. E-mail survey response rates: a review. Journal of Computer Mediated Communication 2001;6:1–20.
Sihvonen T, Lindgren K, Airaksinen O, Manninen H. Movement disturbances of the lumbar spine and abnormal back muscle electromyographic findings in recurrent low back pain. Spine 1997;22:289–95.
Specialist Certification Program Overview Detail-APTA. Available at: http://www.apta.org/Education/specialist/WhyCertify/Overview-SpecCert, accessed November 28, 2004.
Takayanagi K, Takahashi K, Yamagata M, Moriya H, Kitahara H, Tamaki T. Using cineradiography for continuous dynamic-motion analysis of the lumbar spine. Spine 2001;26:1858–65.
Taylor J, O'Sullivan P. Lumbar 'segmental' instability: pathology, diagnosis and conservative management. In: Twomey L, Taylor J, editors. Physical therapy of the low back. 3rd ed. Philadelphia: WB Saunders; 2000. p. 201–47.
Triezenberg J. Teaching ethics in physical therapy education: a Delphi study. Journal Physical Therapy Education 1997;11:16–22.
van Dieen J, Cholewicki J, Radebold A. Trunk muscle recruitment patterns in patients with low back pain enhance the stability of the lumbar spine. Spine 2003;28:834–41.
Vaughan-Williams P, Taylor G, Whittle J. A model framework comparing resources required for activities in the community dental service validated using the Delphi technique. Community Dental Health 1999;16:85–92.
WebSurveyor. 2004 Version 3.6, build 114; WebSurveyor Corporation, 1043 Sterling Road, Suite 204, Herndon, VA, USA.
Werneke M, Hart D. Categorizing patients with occupational low back pain by use of the Quebec Task Force Classification system versus pain pattern classification procedures: discriminant and predictive validity. Physical Therapy 2004;84:243–54.
Who is a Fellow of the American Academy of Orthopaedic Manual Physical Therapists? (FAAOMPT)? Available at: http://www.AAOMPT.org, accessed October 28, 2004.