ORIGINAL ARTICLE

Validity of a single-item measure to assess leg or back pain as the predominant symptom in patients with degenerative disorders of the lumbar spine

A. F. Mannion · U. M. Mutter · T. F. Fekete ·
F. Porchet · D. Jeszenszky · F. S. Kleinstück

Abstract

Purpose Recent studies suggest that the location of predominant pain (back or leg) can be a significant predictor of the outcome of surgery for degenerative spinal disorders. However, others challenge the notion that the predominant symptom can be reliably identified. This study examined the validity of a single item used to determine the most troublesome symptom.

Methods A total of 2,778 patients with degenerative disorders of the lumbar spine scheduled for surgery with the goal of pain relief completed a questionnaire enquiring as to their most troublesome symptom ["main symptom"; back pain (BACK) or leg/buttock pain (LEG)]. They also completed separate 0–10 graphic rating scales for back pain (LBP) and leg/buttock pain (LP) intensity. Receiver operating characteristics (ROC) analysis was used to determine the accuracy with which the "LP minus LBP" score was able to classify patients into their declared "main symptom" group. Sub-studies evaluated the test–retest reliability of the patients' self-rated pain scores ($N = 45$) and the agreement between the main symptom declared by the patient in the questionnaire and that documented by the surgeon after the clinical consultation ($N = 118$).

Results Test–retest reliability of the back and leg pain scores was good ($ICC_{2,1}$ of 0.8 for each), as was patient–surgeon agreement regarding the main symptom (BACK or LEG) (κ value 0.79). In the BACK group, the mean values for pain intensity were 7.3 ± 2.0 (LBP) and 5.2 ± 2.9 (LP); in the LEG group, they were 4.3 ± 2.9 (LBP) and 7.5 ± 1.9 (LP). The area under the curve for the ROC was 0.95 (95 % CI 0.94–0.95), indicating excellent discrimination between the BACK and LEG groups based on the "LP minus LBP" scores. A cutoff score >0.0 for "LP minus LBP" score gave optimal sensitivity and specificity for indicating membership of the LEG group (sensitivity 79.1 %, specificity 95.7 %).

Conclusions The responses on the single item for the "main symptom" were in good agreement with the differential ratings on the 0–10 pain scales for LBP and LP intensity. The cutoff >0 for "LP minus LBP" for classifying patients as LEG pain predominant seemed appropriate and suggests good concurrent validity for the single-item measure. The single item may be of use in sub-grouping patients with the same disorder (e.g. spondylolisthesis) or as an indication in surgical decision-making.

Introduction

The pain associated with many degenerative disorders of the lumbar spine can be exclusively located in the lower back, referred to the legs/buttock region, or manifest as a combination of both. It has been suggested that, since surgical treatment is only appropriate when the specific pain generator is known [1], determination of the aetiology of the pain is required to dictate both the need (or otherwise) for surgery [2] and the primary surgical procedure to be applied [1]. It has been a long-held belief in clinical practice that leg pain dominant over back pain is an important factor in determining the degree of surgical success, especially in relation to

A. F. Mannion (✉) · U. M. Mutter · T. F. Fekete · F. Porchet ·
D. Jeszenszky · F. S. Kleinstück
Spine Center Division, Department of Research
and Development, Schulthess Klinik, Lengghalde 2,
8008 Zurich, Switzerland
e-mail: anne.mannion@yahoo.com

decompression surgery [1, 3, 4]. However, until recently, this empirical belief had not been formally tested in studies that had actually quantified relative leg pain and back pain intensities. This situation was redressed in two recent studies in which it was verified that the patient's declared predominant pain location (leg or back) and the difference in back and leg pain intensities recorded on a pain scale were indeed significant predictors of the outcome of decompressive surgery for spinal stenosis [5] and herniated disc [6]. Its role in predicting the outcome of degenerative spondylolisthesis remains controversial, with some authors reporting a significant relationship [7] and others failing to observe any significant effect [8]. However, the validity of these findings is threatened somewhat by studies that challenge the reliability and validity of methods used for determining "leg-dominant pain" [1]. Cross-validation between different instruments purporting to measure the same phenomenon (i.e., dominant pain location) revealed weighted κ values of 0.27–0.65, depending on the instruments compared.

This aim of this study was to examine the concurrent validity of a single-item measure, commonly used for determining the most troublesome symptom [5, 6] and included in the Eurospine Spine Tango patient-outcome form [9], by comparing it with simultaneous measures of back pain and leg pain intensity measured on 0–10 pain scales. In sub-studies, the test–retest reliability of the pain scales themselves, and the agreement between the predominant pain location declared by the patient in the questionnaire and documented by the surgeon after clinical interview were also analysed.

Methods

Patients

The study was a retrospective analysis of prospectively collected data, carried out using the framework of the Spine Society of Europe (SSE) Spine Tango Spine Surgery Registry together with our own local spine surgery outcomes database. It included the data from consecutive patients with degenerative disorders of the lumbar spine who had undergone spine surgery for pain relief at our own Spine Center, from March 2004 to Feb 2011. The patients had to have a good understanding of written German or English or (after 2006) French, Spanish, Italian or Portuguese, and satisfy the study's surgical admission criteria. The latter made use of the options ticked in relation to the given fields on the Spine Tango surgery form and were as follows: operation on the lumbar or lumbosacral region of the spine; degenerative disease as the main pathology; goal of surgery includes pain relief.

A total of 2,778 eligible patients were identified [1,437 women, 1,341 men; mean (SD) age 61.4 years].

Questionnaires

Before surgery, patients were requested to complete the multidimensional Core Outcome Measures Index (COMI) questionnaire [10, 11]. The questionnaire was sent to the patient at home, along with the information about their forthcoming hospital stay, and they were asked to complete it and hand it in during admission. Completion of the questionnaire at home ensured that the information provided by the patient was free of any care provider's influence. The COMI is a multidimensional index consisting of validated questions covering the domains of pain [back pain (LBP) and leg/buttock (sciatic) pain (LP) intensity, each measured separately on a 0–10 Graphic Rating Scale], function, symptom specific well-being, general quality of life, and social and work disability. (Only the data concerning the pain scales were relevant to the present study.) Patients also responded to a multiple-choice question enquiring: "which of the following problems troubles you the most?", with response options of back pain, leg/buttock (sciatic) pain, neurological disturbances, none of the above. They were requested to tick one answer only, and only the data of patients indicating back pain (BACK) or leg pain (LEG) as the main problem were considered further in the present study. The whole COMI questionnaire was preceded by a short introduction explaining that "back problems can lead to various symptoms such as back pain and/or pain in the legs/buttocks, as well as to neurological disturbances such as tingling, 'pins and needles', or numbness in any of these regions". This was considered to orientate and direct the patients appropriately regarding "back-related" leg pain as opposed to any other general leg pain that they may be experiencing.

To assess test–retest reliability of the individual pain scales, the data collected (but as yet unpublished) from a previous study to validate the individual items contained in the COMI [10] were analysed. In brief, 45 patients with chronic LBP completed the questionnaires twice over a period of 1–2 weeks (questionnaires sent out twice by mail).

To assess the concurrent validity of the "main symptom" item, a further sub-study of 118 consecutive patients and their treating surgeons ($N = 2$) was carried out to compare the agreement between patient and surgeon for their response to this item. The patients were all attending the surgeon's practice for consultation in relation to a problem of the lumbar spine (of any sort). The answer to the single-item question was documented by the surgeon during the preoperative consultation; after the consultation, the patient was requested to independently complete a questionnaire containing the single-item measure and the pain scales (as part of the COMI), and to hand it in to the secretary before leaving the hospital.

Statistical analyses

Descriptive data are presented as means ± standard deviations (SD). The significance of the difference in mean values for the two completions of the pain scales was analysed using one-way repeated measures ANOVA, and from this, the intraclass correlation coefficient ($ICC_{2,1}$) for the repeated trials was determined, to assess test–retest reliability.

Agreement regarding the "main problem" declared by the patient in the questionnaire and that documented by the surgeon during the consultation was assessed as the percentage absolute agreement and using κ values.

A "LP minus LBP score" was calculated from the pain scores given on the two 0–10 scales. The sensitivity and specificity of individual LP-minus-LBP scores in "predicting" the main symptom (LEG as opposed to BACK) was examined using the Receiver Operating Characteristic (ROC) method. This can be considered analogous to evaluating the accuracy of a diagnostic test, in which the "LP–LBP score" is the diagnostic test and the "main symptom" represents the gold standard [12]. The ROC curve synthesises information on sensitivity and specificity for predicting the "main symptom" and consists of a plot of 'true-positive rate' (sensitivity) versus 'false positive rate' (1-specificity) for each of a range of possible LP–LBP scores. Thus, sensitivity and specificity are calculated for a LP–LBP score of −1.0 points, −0.5 points, 0 points, 0.5 points, 1.0 points, and so on. The area under the ROC curve (AUC) was interpreted as the probability of correctly discriminating between patients with the main symptom being "LEG" (as opposed to BACK) based on the LP–LBP score. The AUC can range from 0.5 (no accuracy in discriminating) to 1.0 (perfect accuracy in discriminating).

In summary, the ROC analysis was used to: (a) determine how well the "LP minus LBP score" was able to correctly classify patients into their declared "main symptom" group (LEG or BACK); and (b) to identify the cutoff value for the "LP minus LBP" score that most accurately classified a patient as belonging to the LEG group.

The analyses were conducted using Statview 5.0 (SAS Institute Inc, San Francisco, CA, USA) and SPSS (version 16, SPSS Inc, Chicago, IL, USA) and statistical significance was accepted at the $p < 0.05$ level.

Results

Pain scores in relation to "declared main symptom"

Thousand five hundred and seventy five (56.7 %) patients were in the LEG group and 1,203 (43.3 %) in the BACK group. The individual values for the "LP minus LBP" scores in each of the two groups are shown in Fig. 1. In
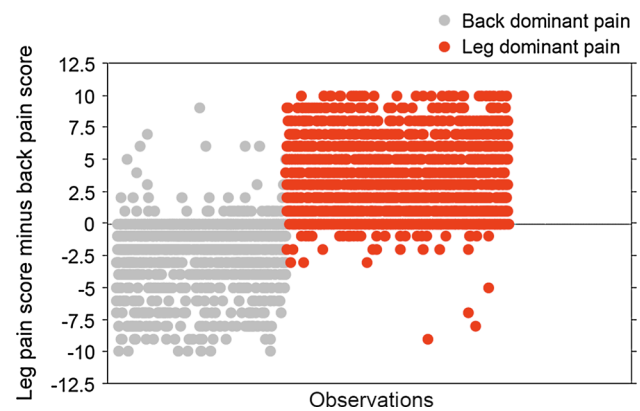

Fig. 1 Distribution of individual "leg pain minus back pain" scores for individuals who declared that they had either back-dominant or leg-dominant pain

total, 619 (22.3 %) patients reported equal leg and back pain (LP minus LBP = 0), but were still able to indicate one problem as being more troublesome than the other in the "main symptom" question.

In the LEG group, the mean value for LP intensity was 7.5 ± 1.9 and for LBP, 4.3 ± 2.9 (LBP); the corresponding values for the BACK group were 5.2 ± 2.9 (LP) and 7.3 ± 2.0 (LBP). The mean values for the LP scores and for the LBP scores were significantly different between the LEG and BACK groups (each, $p < 0.0001$).
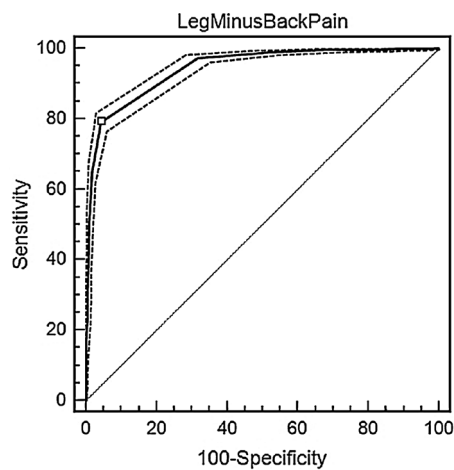
The AUC for the ROC (Fig. 2) was 0.945 (95 % CI 0.935–0.953), which indicated excellent discrimination of patients into the LEG as opposed to the BACK group based on their "LP minus LBP" scores. A cutoff score >0.0 for the "LP minus LBP score" on the pain scales gave optimal sensitivity and specificity for indicating membership of the LEG rather than BACK group (sensitivity 79.1 %, specificity 95.7 %).

Test–retest reliability

One of the 45 patients in the reliability study failed to complete the leg pain scale (at both the first and second assessment), and another failed to complete the pain scales at the first assessment (it appeared that he had overlooked that page of items in the questionnaire booklet).

For the remaining patients, the mean values for pain intensity measured on the 0–10 pain scales at the two time-points of assessment were: LP (N = 43) 4.0 ± 2.6 (first trial) and 4.2 ± 3.0 (second trial) ($p = 0.46$) and LBP (N = 44) 5.4 ± 2.6 (first trial) and 5.8 ± 2.6 (second trial) ($p = 0.15$). The test–retest reliability of the pain scales was good, with intraclass correlation coefficients ($ICC_{2,1}$) of 0.79 for LP and 0.77 for LBP.

The mean values for the difference between LP and LBP scores (LP minus LBP) at the two time-points were

Fig. 2 Receiver operating characteristics (ROC) curve for the discrimination of patients into LEG-predominant as opposed to BACK-predominant based on their "LP minus LBP" scores (*upper panel*); sensitivity and specificity for various cutoff criteria (*lower panel*). See text for details

-1.3 ± 2.7 (first trial) and -1.5 ± 2.7 (second trial) ($p = 0.61$), with an $ICC_{2,1}$ of 0.62.

Agreement regarding the "main problem" declared by the patient and surgeon

In the group of 118 patients, 58 (49 %) declared LBP to be their main problem and 51 (43 %) leg pain. 9 patients (8 %) declared neurological disturbance to be their main problem (consecutive lumbar spine patients were included in this sub-study, and so, unlike in the main study, not all were consulting with pain as their main problem).

For the 58 cases where LBP was the main problem declared by the patient, the doctor's assessment during the consultation was in agreement in 54 of them; in three cases the doctor indicated LP and in one case, neurological disturbances as the main problem. For the 51 cases where the patient reported LP to be main problem, the doctor agreed in 46; in three cases the doctor indicated LBP and in two cases, neurological disturbances. For the nine cases where the patient reported neurological disturbance to be the main problem, the doctor agreed in four; for the other five cases, he/she had indicated leg pain as the main problem.

Overall, the proportion of agreement for the patients' and doctors' ratings was 88 %, and the κ value was 0.79 (95 % CI 0.69–0.89), which is considered "good".

Discussion

The aim of the present study was to assess the validity of a single item used for assessing the "most troublesome symptom" in patients with painful degenerative disorders of the lumbar spine. We found that the responses on the single item were in good agreement with the corresponding relative scores for back pain and leg pain intensity indicated on 0–10 graphic rating pain scales. The area under the curve for the ROC analysis (AUC, 0.95) indicated excellent discrimination of patients into LEG-predominant as opposed to BACK-predominant based on their "LP minus LBP" scores. The cutoff >0 for "LP minus LBP score" for classifying patients as belonging to the LEG group seemed appropriate and suggested good concurrent validity for the item. The test–retest reliability of the pain scales themselves was also good (i.e. greater than 0.7 [13]), as was the agreement between the patient's worst problem as self-declared in the questionnaire and that recorded by the surgeon during the preceding clinical assessment and history-taking (κ = 0.78). The $ICC_{2,1}$ was a little lower for the "LP–LBP score difference" (0.62) than for either of the LBP or LP absolute scores (approximately 0.8), despite its being calculated from these absolute pain scores. This was likely because the range of values for the score difference was not as great as for either of the individual pain scales, and increased homogeneity within a group of values typically reduces the ICC [14].

Although the "predominant pain location" is (anecdotally) frequently used in guiding treatment, few studies have sought to assess the reliability or validity of its measurement. Wai et al. [1] used 8 self-administered instruments to assess the predominant pain location and, depending on the instrument used, found that up to 32 % patients provided a completely opposite response on retest compared with their initial response. When the same instruments were administered by an interviewer, the figure was <10 %. This may

have indicated that in the self-administered versions the clarity of the wording was not optimal and that problems of comprehension may have led to the poor reliability for some instruments. This is also suggested by the finding that their self-administered instrument suffered from a higher proportion of missing data, with up to 25 % patients failing to answer an item or providing an inappropriate response on either the test or retest [1]. In the present study, only one patient out of 45 in the reliability study failed to complete the leg pain scale (and did so on both test and retest, such that we presume the patient simply had no leg pain at all and hence chose not to mark anything), and in one patient the pain scales were not completed pre-operatively (it appeared that the patient had overlooked the whole page containing these items in the questionnaire booklet). Possibly, the method of questioning applied in the present study and the introductory paragraph [explaining how the symptoms of spinal problems could manifest themselves in different ways e.g. as back pain, leg/buttock (sciatic) pain, neurological disturbances] were easier to understand, leading to less missing data and good reliability.

In the study of Wai et al. [1], cross-validation between different instruments purporting to measure the same phenomenon (dominant pain location) revealed weighted κ values of 0.27–0.65. When the results of all the psychometric quality tests were considered together, their "percent back, percent leg pain" item appeared to perform best of all, followed by their Numeric Rating Scale (NRS), and a question on the region as to which the patient would assign priority for treatment. Interestingly, although the main conclusion that appeared to emerge from the study of Wai et al. [1] was that there was a lack of reliability for many of the instruments used, the authors did concede in their discussion that patients *are* able to reliably assess their pain dominance location if asked with a specific question, in a structured format, using questions that force them to quantitatively evaluate their back and leg pain. Indeed, in a further study by the same research group, the answers given in response to all three "best-performing" items (percent, NRS and treatment items) were together found to allow effective triaging of surgical candidates [2]. Although three questions is not a great deal, our current philosophy for obtaining good patient-compliance with outcome measurement in our hospital is to keep our questionnaires as short and succinct as possible [11]. Hence, we use just one question for directly assessing the predominant problem—"which problem troubles you the most", which is not dissimilar to Wai et al's [1] treatment question "which problem would you preferentially direct treatment at". The disadvantage of using *only* a "% of each" or "predominance" question is that one does not have an individual score for absolute back pain and leg pain severity and these two variables are often of interest in themselves, especially in research studies of treatment outcome. Hence, to allow quantitative analyses of relative pain, we would recommend assessing the pain intensity for back and leg separately, too, rather than only asking the predominance question.

Given that the most troublesome symptom can theoretically be derived from the information in the pain scales, one may question whether the "main problem" question is actually required. However, there are occasions when an individual rates both regions as similarly painful (this occurred in approximately one-fifth of cases in the present study; see "Results"), yet is still able to acknowledge one problem as being more troublesome/bothersome than the other. It is also possible that the patient's main problem is not the one with the greatest pain intensity (as suggested by the ROC area under the curve of 0.95 rather than 1.00, and the less than 100 % sensitivity and specificity for the LP–LBP scores in predicting group membership), which would render the item useful in its own right as a measure of symptom "bothersomeness". Further, for descriptive statistics or statistical analyses, an item that is dichotomous at source can sometimes provide greater clarity and facilitate interpretation of the data compared with continuous scores (such as pain ratings). Further, in addition to pain, one may wish to add other options to the "main problem" item, such as motor function, deformity or body image, allowing the reason for undergoing treatment in *non*-painful complaints (where scores on pain scales may appear incongruous with measures of function, quality of life, etc.) to be ascertained. Finally, the "main/greatest problem" item represents a simple tool that can be used in clinical practice in the absence of formal pain scales, e.g. during the clinical consultation, to provide a valid appraisal of the most troublesome symptom and guide treatment.

A limitation of the present study concerns the uncertainty regarding the definition of buttock pain and the location to which it is best classified as belonging, back or leg. It has been argued that, in spinal disorders that have a component of neural compression, most clinicians would agree that the combination of buttock pain and leg pain represents the referred pain, differentiating it from the more mechanical low back pain [5]. However, patients often report pain in complex patterns, with overlap between different regions [15], and it may be hard for the patient to differentiate between back pain and buttock pain, since there is no clear boundary between these anatomical sites. The literature does not offer a clear definition of what constitutes buttock pain and what constitutes back pain, leaving the clinician and the patient to make the final assessment, as they did in the present study.

In conclusion, the results of this study suggest good concurrent validity for the single-item measure of the "main symptom". It may be of use in future studies for

sub-grouping patients with the same disorder (e.g. in spondylolisthesis) or as an indication to guide surgical decision-making.

Conflict of interest None.

References

1. Wai EK, Howse K, Pollock JW, Dornan H, Vexler L, Dagenais S (2009) The reliability of determining "leg dominant pain". Spine J 9:447–453
2. Simon D, Coyle M, Dagenais S, O'Neil J, Wai EK (2009) Potential triaging of referrals for lumbar spinal surgery consultation: a comparison of referral accuracy from pain specialists, findings from advanced imaging and a 3-item questionnaire. Can J Surg 52:473–480
3. Postacchini F (1999) Surgical management of lumbar spinal stenosis. Spine 24:1043–1047
4. Atlas SJ, Keller RB, Robson D, Deyo RA, Singer DE (2000) Surgical and nonsurgical management of lumbar spinal stenosis: four-year outcomes from the maine lumbar spine study. Spine 25:556–562
5. Kleinstuck FS, Grob D, Lattig F, Bartanusz V, Porchet F, Jeszenszky D, O'Riordan D, Mannion AF (2009) The influence of preoperative back pain on the outcome of lumbar decompression surgery. Spine 34:1198–1203
6. Kleinstueck FS, Fekete T, Jeszenszky D, Mannion AF, Grob D, Lattig F, Mutter U, Porchet F (2011) The outcome of decompression surgery for lumbar herniated disc is influenced by the level of concomitant preoperative low back pain. Eur Spine J 20:1166–1173
7. Pearson A, Blood E, Lurie J, Abdu W, Sengupta D, Frymoyer JW, Weinstein J (2011) Predominant leg pain is associated with better surgical outcomes in degenerative spondylolisthesis and spinal stenosis: results from the Spine Patient Outcomes Research Trial (SPORT). Spine (Phila Pa 1976) 36:219–229
8. Kleinstueck FS, Fekete TF, Mannion AF, Grob D, Porchet F, Mutter U, Jeszenszky D (2012) To fuse or not to fuse in lumbar degenerative spondylolisthesis: do baseline symptoms help provide the answer? Eur Spine J 21:268–275
9. Zweig T, Mannion AF, Grob D, Melloh M, Munting E, Tuschel A, Aebi M, Roder C (2009) How to Tango: a manual for implementing Spine Tango. Eur Spine J 18(Suppl 3):312–320
10. Mannion AF, Elfering A, Staerkle R, Junge A, Grob D, Semmer NK, Jacobshagen N, Dvorak J, Boos N (2005) Outcome assessment in low back pain: how low can you go? Eur Spine J 14:1014–1026
11. Mannion AF, Porchet F, Kleinstück F, Lattig F, Jeszenszky D, Bartanusz V, Dvorak J, Grob D (2009) The quality of spine surgery from the patient's perspective: part 1. The Core Outcome Measures Index (COMI) in clinical practice. Eur Spine J 18:367–373
12. Deyo RA, Centor RM (1986) Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Dis 39:897–906
13. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60:34–42
14. Hopkins WG (2000) Measures of Reliability in Sports Medicine and Science. Sports Med 30:1–15
15. Davies HT, Crombie IK, Macrae WA (1998) Where does it hurt? Describing the body locations of chronic pain. Eur J Pain 2:69–80