European Journal of Physical and Rehabilitation Medicine
EDIZIONI MINERVA MEDICA

Responsiveness and minimal important change of the Quebec Back Pain Disability Scale in Italian patients with chronic low back pain undergoing multidisciplinary rehabilitation.

Marco MONTICONE, FEDERICO ARIPPA, Calogero FOTI, Franco FRANCHIGNONI

European Journal of Physical and Rehabilitation Medicine 2022 Feb 01
DOI: 10.23736/S1973-9087.22.07385-3

Article type: Original Article

Article first published online: February 1, 2022
Manuscript accepted: January 18, 2022
Manuscript received: November 29, 2021

Responsiveness and minimal important change of the Quebec Back Pain Disability Scale in Italian patients with chronic low back pain undergoing multidisciplinary rehabilitation.

Running Title: QBPDS responsiveness and MIC

Authors: Marco Monticone[1,2], Federico Arippa[2], Calogero Foti[3], Franco Franchignoni[4]

Affiliations

[1] Dept. Medical Sciences and Public Health, University of Cagliari, Cagliari, Italy; Neurorehabilitation Unit, Dept. Neuroscience and Rehabilitation, G. Brotzu Hospital, Cagliari, Italy.

[2] Neurorehabilitation Unit, Dept. Neuroscience and Rehabilitation, G. Brotzu Hospital, Cagliari, Italy.

[3] Physical and Rehabilitation Medicine, Dept. Clinical Sciences and Translational Medicine, University of Rome Tor Vergata, Rome, Italy.

[4] Istituti Clinici Scientifici Maugeri, PRM Unit of Tradate Institute, Tradate (VA), Italy

Name and address for correspondence

Marco Monticone, MD, PhD, Prof.

Cittadella Universitaria, Strada Statale, 554 - Monserrato, Cagliari, Italy

Phone: +39 0706753109; Fax: +39 07051096487

Email: marco.monticone@unica.it

ABSTRACT

Background. There is still a lack of information concerning MIC of the QBPDS, that limits its use for clinical and research purposes.

Aim. Evaluating responsiveness and minimal important change (MIC) of the Quebec Back Pain Disability Scale (QBPDS) in Italians with chronic low back pain (LBP).

Design. Methodological research based on an observational study.

Setting. Outpatient rehabilitation hospital.

Population. Two hundred and one patients with chronic LBP.

Methods. At the beginning and end of a multidisciplinary rehabilitation programme, patients completed the QBPDS. At the end of treatment, they completed a 7-level global perceived effect (GPE) scale, which was split to obtain a dichotomous outcome (improved vs. stable). Responsiveness was calculated by distribution-based [effect size (ES); standardised response mean (SRM); minimum detectable change (MDC_{95})] and anchor-based methods [Receiver Operating Characteristics (ROC) curves]. ROC curves were also used to compute the MIC (based on QBPDS change score, both absolute and expressed as percentage). Correlations between the change score of the QBPDS and GPE were calculated.

Results. The ES was 0.29, the SRM was 0.43, and the MDC_{95} was 12 points. ROC analysis of the absolute change scores showed a MIC value of 6 points, with an area under the curve (AUC), sensitivity, and specificity of 0.83 (95%C.I. 0.77-0.90), 77.7% and 80.8%, respectively. ROC analysis based on the percent change score from baseline revealed a MIC of 18% with an AUC, sensitivity and specificity of 0.85 (95%C.I. 0.79-0.91), 80.6% and 80.8%, respectively. Correlation between change score of the QBPDS and GPE was ρ=-0.67.

Conclusions. The QBPDS score change (expressed in both absolute value and percentage from baseline) was sensitive in detecting clinical changes in Italian subjects with chronic LBP undergoing multidisciplinary rehabilitation. In clinical practice, we recommend –where absolute change is lower than MDC– to rely on the MIC taking into account the percentage change from baseline condition.

Clinical Rehabilitation Impact. The present study investigated the responsiveness and MIC of the QBPDS in a group of patients with chronic LBP. Our findings showed that the QBPDS score may classify with good to excellent discriminatory accuracy subjects who consider themselves as improved. Where examining change, we recommend to consider both MICs we provided (expressing score change both in absolute value and as a percentage from baseline), and disregard values lower than MDC_{95}, not being discernible from measurement error.

Key words

Low back pain; Rehabilitation; Responsiveness; Minimal Important Change; Quebec Back Pain Disability Scale.

INTRODUCTION

Responsiveness and minimal important change (MIC) represent critical issues to research and patient management in measuring clinical change. The first corresponds to the ability of an instrument to detect changes in the construct to be measured over time while the second represents the smallest amount of change in score of the construct to be measured that patients perceive to be significant [1,2]. Researchers need such information to state power calculations, sample size estimates and cost evaluations of given interventions, as well as for evaluating overall prognosis of diseases. Further, these properties are important in clinical care when assessing interventions effectiveness and guiding clinicians' decision making [1].

To the Authors' knowledge, responsiveness of the Quebec Back Pain Disability Scale (QBPDS) was calculated in patients with acute and chronic LBP in English, Dutch, French and Portuguese populations [3–10]. Figures on MIC of the above scale were provided in patients with acute and chronic LBP in Dutch and Portuguese samples [7,8,10]. Recently, a systematic review on psychometric properties was conducted and, based on the methodological drawbacks found, the Authors stated more high-quality, context- and population-specific studies are needed concerning both responsiveness and MIC, to be conducted in all existing language versions [11].

The Italian version of the QBPDS has been psychometrically analysed in patients with chronic LBP [12]. Factor analysis suggested a one-factor 20-item structure (first factor variance explained of 54.7%), and the tool was found to have similar properties to those of the original version [3]: internal consistency of 0.95, test-retest reliability of 0.90, and good construct validity, showing moderate correlations with the Roland Morris Disability Questionnaire (r=0.40), the Oswestry Disability Index (r=0.48), and a pain numerical rating scale (r=0.44) [12]. However, its responsiveness and MIC have not yet been determined, therefore limiting its use for research and clinical purposes.

The aim of this study was therefore to determine the responsiveness and MIC of the QBPDS in Italian patients with chronic LBP undergoing multidisciplinary rehabilitation using both distribution-based and anchor-based methods suggested in the current literature and based on the ‘‘COnsensus-based Standards for the selection of health status Measurement INstruments’’ (COSMIN) [2,13,14].

METHODS

Patients

Outpatients consecutively admitted to Physical and Rehabilitation Units at the University Hospital in XX (XX), and at the University Hospital in XX (XX) were enrolled between September 2016 and July 2018 [12]. The inclusion criteria were: a diagnosis of chronic non-specific LBP (i.e. documented history of pain lasting for more than twelve weeks) [15], good understanding of Italian language, and an age of >18 years. The exclusion criteria were: ì) acute (lasting up to four weeks) and subacute non-specific LBP (lasting up to twelve weeks); ìì) specific causes of LBP (e.g. disc herniation, canal stenosis, spinal deformity, fracture, spondylolisthesis, or infections), central or peripheral neurological signs, ruled out by case history and/or imaging (lumbar radiographs and, in doubtful cases, Computed Tomography or Magnetic Resonance Imaging), systemic illness, cognitive impairment (Mini Mental State Examination of <24), recent myocardial infarctions, and cerebrovascular events; ììì) previously prescribed physical or cognitive-behavioural therapy; and ìv) refusal or inability to adhere to the treatment.

Patients' sociodemographic and clinical characteristics were investigated using a specific form including information on: age, gender, body mass index, pain duration, education level, occupation, and marital status.

This research was part of an observational study approved by the Institutional Review Board of the local Hospital (date of approval: 07/04/2016, n° 12). All patients gave their written consent to participate and the study was conducted in conformity with ethical and humane principles of research.

Outcome measures

THE QUEBEC BACK PAIN DISABILITY SCALE (QBPDS) – This questionnaire includes 20 items and refers primarily to the main daily living activities that are frequently affected by LBP. These items allow patients to rate their degree of restriction in activities from 0 ('not difficult at all') to 4 ('unable to do'). The responses to the items are added and total score ranges from 0 to 100, with higher scores indicating greater disability [3,16].

THE GLOBAL PERCEIVED EFFECT – Patients' global perceived effect of the intervention associated to daily living activities related to LBP (GPE) was evaluated at the end of treatment using the following question: "*Overall, how much did the treatment you received help your daily activities due to current LBP?*". The perceived effect was scored using a 7-level Likert scale with following response options: -3 "made things a lot worse"; -2 "made things worse"; -1 "made things partially worse"; 0 "did not help (unchanged)"; +1 "helped only a little"; + 2 "helped"; + 3 "helped a lot" [17].

Procedure

Participants underwent an individual 60-min motor training sessions twice a week for an eight-week outpatient programme, that included exercises aimed at improving postural control, strengthening and stabilising the back muscles, and stretching. The programme was the same for all of the patients and was previously tested for efficacy by means of a randomized controlled pilot study [18].

All patients were required informed consent concerning research aim, questionnaires and procedures, before study began. Questionnaires were administered as part of the pre/post-rehabilitation assessment by secretarial staff, who checked them and returned any uncompleted part to the participants for completion, in order to minimise the rate of missing/multiple responses.

Mild analgesics and non-steroidal anti-inflammatory drugs (NSAIDs) were allowed during the study. No more than one tablet per day on demand for more than 7-10 successive days was permitted and drugs intake, and patients' symptoms and needs were regularly supervised.

Statistical analysis

Indexes of reliability and responsiveness were calculated, being reliability a critical component of responsiveness.

Reliability has been determined as: ì) internal consistency, calculating Cronbach's α; the closer to 1, the higher the internal consistency of the scale's items. Alpha values of >0.70 are recommended for group-level comparisons, whereas a minimum of 0.85 to 0.90 is desirable for individual judgments; ìì) test-retest reliability of global scores, by means of the Intraclass Correlation Coefficient, with a "two-way mixed effects, single measurement" model ($ICC_{2.1}$) [2].

Responsiveness was investigated by means of distribution- and anchor-based methods. Regarding the former, which analyses the ability to detect change in general, we calculated [19]:

1) the effect size (ES), representing the difference between the pre- and post-test scores divided by the pre-test standard deviation (SD) and classified as trivial (when figures are <0.2), small (0.2-0.39), moderate (0.4-0.7), or large (>0.7);

2) the standardized response mean (SRM), representing the ratio between individual change and the SD of that change, with figures of 0.20, 0.50, and 0.80 respectively showing small, moderate, and large changes;

3) the standard error of measurement (SEM) based on the following equation, with the $ICC_{2.1}$ value taken from test-retest results of a previous reliability study on the same sample [12]: $SEM = SD\sqrt{1 - ICC_{2,1}}$;

4] the minimum detectable change [MDC] which is based on SEM and calculated as in the following equation: $MDC = SEM * z\ value * \sqrt{2}$.

The MDC represents the smallest change in score that probably reflects true change, and not simply measurement error. Its 95% confidence level (MDC_{95}) is calculated with a *z* value of 1.96 in the above equation [20].

The latter methods include GPE as an external criterion (anchor) to determine whether changes in outcome scores were clinically meaningful [13,14,21]. We analysed two parameters:

1) for the mean change approach, we calculated the mean change of participants graded on the GPE as not improved (GPE= 0), minimally improved (GPE= 1), moderately improved (GPE= 2), and largely improved (GPE= 3);

2) for the receiver operating characteristic (ROC) curve approach, we determined the best cut-off score and the area under the curve (AUC) after splitting the participants into two groups: stable (GPE= 0 or 1) and improved (GPE= 2 or 3).The AUC indicates the probability of correctly identifying, in randomly selected pairs of patients who have and have not improved, the one who has improved. The greater the AUC, the greater a measure's ability to distinguish those with versus without a meaningful improvement, with an AUC >0.70 indicating an acceptable discrimination and >0.80 a good one. The cut-off points (MIC) based on QBPDS change score (both absolute and expressed as percentage) from baseline were estimated by considering the intersection with the 45-degree line (i.e. antidiagonal line) which indicates the point that jointly maximized sensitivity and specificity (being associated with the least amount of misclassification) [22].

Further, we studied GPE correlation with the pre-post treatment change score in order to check its validity as a criterion anchor, hypothesizing to find a nontrivial correlation (>0.30-0.50) [23].

The analyses were made using IBM SPSS v.26 software (IBM Corp., Armonk, NY).

Data availability

The data associated with the paper are not publicly available but are available from the corresponding author on reasonable request.

RESULTS

Of the 260 patients invited to participate, 11 did not meet the inclusion criteria (specific causes of LBP, n=5; systemic illness, n=4; cognitive impairment, n=2), 35 refused, and 13 were unable to adhere to the treatment (logistic problems, n=8; economic difficulties, n=2; personal problems, n=3).

The final study population consisted of 201 patients with a mean age of 48.2±11.8 years and a median pain duration of 56 months (range 6–80). At baseline, average pain intensity value was 4.0±1.3 (on a 0-10 numerical rating scale) and the average lumbar disability 23.7±8.0 points (on the 0-100 Oswestry Disability Index). The body mass index was 24.3±3.7 kg/m^2. Table I shows socio-demographic characteristics of the patients.

Please insert Table I approximately here

Pain was treated through symptomatic drugs: 34 subjects (17%) needed paracetamol, and 20 subjects (10%) NSAIDs. No patient took more than one tablet per day on demand for more than 7-10 consecutive days.

Table II describes baseline and post-treatment scores of the QBPDS-I scale. The QBPDS scores did not significantly differ between males and females (baseline: 32.7±14.7 vs. 31.8±12.1; post-treatment 27.9±13.9 vs. 28.5±13.8) or between younger and older participants, split according to the median age of the sample, 48 years (baseline: 33.1±12.1 vs. 31.3±14.2; post- treatment 29.0±14.5 vs. 27.6±13.2).

Please insert Table II approximately here

Cronbach's α and test-retest reliability ($ICC_{2.1}$) [12] indexes for QBPDS are shown in Table III, along with the related distribution-based indices for responsiveness.

Please insert Table III approximately here

Following the dichotomisation on the basis of the GPE score, 103 participants (51%) were classified as improved (GPE=2 or 3) and 78 (39%) as stable (GPE=0 or 1). The remaining 20 (10%) showed a worsened clinical condition according to their GPE change score (<0), and were excluded from further analyses. As there were more than 50 subjects per subgroup, these estimates assured an adequate sample size for calculating responsiveness [24]. The correlation of the score change of the QBPDS with the GPE was ρ= -0.67, while that with baseline and final QBPDS score was ρ= 0.22 and ρ= -0.35, respectively.

Distribution-based methods showed low to moderate effects for QBPDS score. The ES, and the SRM were 0.29, and 0.43, respectively.

For anchor-based methods, the mean QBPDS raw and percent score changes according to the GPE ratings were reported in Table IV.

Please insert Table IV approximately here

ROC analysis of the absolute change scores in QBPDS revealed an AUC of 0.83 (95% C.I. 0.77-0.90), showing a high ability in identifying between improved and stable patients (Figure 1). The optimal cut-off (i.e. MIC), chosen by considering the intersection with the antidiagonal line was 6 points, which was associated with 77.7% sensitivity, 80.8% specificity and 78.1% accuracy.

Please insert Figure 1and 2 approximately here

The additional ROC analysis based on the percent change score from baseline revealed an optimal cut-off point of 18% (AUC=0.85, 95% C.I. 0.79-0.91), which was associated with 80.6% sensitivity, 80.8% specificity and 79.5% accuracy.

DISCUSSION

This is the first study to examine the responsiveness and the MIC of the QBPDS in Italian patients with chronic LBP undergoing multidisciplinary rehabilitation, by means of distribution- and anchor-based methods.

The major concern when referring to distribution-based methods is that they are related to the magnitude of change scores but cannot provide a meaningful indication about the clinical importance of the observed change [25–27]. On the contrary, the accuracy of the anchor-based methods depends on many variables, such as baseline values, type of population, characteristics of the context, choice of the anchor, or the cut-off used [2,23]. Basically, a good way to determine MIC is to compare and interpret the information gathered by multiple reference standards, calculated on the same sample according to the above methods [25,28].

We used distribution-based methods to examine some statistical group-level characteristics of scores in our sample. The ES and SRM showed low to moderate responsiveness of the QBPDS to the multidisciplinary rehabilitation programme. When other studies are taken into account, Kopec et al. [3] reported similar values to those presented in this manuscript, while Wilhelm and colleagues [9] found slightly higher values for ES. SRM results here obtained are in accordance to Davidson et al.'s study [5], and lower of those reported by Mens and colleagues [6] and Wilhelm et al. [9].The MDC_{95} in our study [about 12 points] was lower than those reported in previous studies [4,5,7,8,10], which ranged from 15.8 [8] to 32.9 points [7], but still higher than the MIC related to absolute change scores.

As for anchor-based methods, the good correlation found between changes in QBPDS score and GPE indicated that QBPDS is high reflective of subjects' perception about their change in daily living activities during the treatment period. As expected, the QBPDS score significantly decreased as the GPE score increased, meaning an improvement in patients' health status.

ROC analysis showed the good capability of QBPDS absolute change to classify subjects as improved or not, based on their GPE (Figure 1). When considering the difference between baseline and follow-up scores, a change of at least 6 points represented the optimal cut-off value, identifying in our sample a clinically important change for an individual. Other authors also investigating chronic complaints identified MICs similar to our estimates (Dutch study: 5 points;

Portuguese study: 6.5 points) [8,10]. Despite the relative consistency of these results, a MIC value lower than the MDC is problematic, because the MDC should be interpreted as a preliminary step toward establishing a (higher) MIC, by benchmarking it to the boundaries of measurement error. In practice, a change higher than this MIC (6 points) but lower than the MDC (about 12 points) represents a clinical improvement with a considerable chance of being due to the measurement error, and thus without a straightforward clinical interpretation.

On the other hand, when taking into account the percent score change with respect to the QBPDS baseline value, the ROC analysis revealed an optimal cut-off of at least 18%. This value is also in line with those proposed by previous studies of 18-24% [8,10], and just slightly lower to the parsimonious value proposed by an expert panel (20-30%) as a preliminary guidance, based on a limited and heterogeneous empirical evidence (difficult to integrate) and clinical judgment [29].

In light of such results, we consider that higher importance should be given to the change scores expressed as percentage change from baseline, when assessing the efficacy of a multidisciplinary treatment in patients with chronic LBP [8, 30]. Moreover, as already acknowledged [29], different MICs should be calculated and used in different contexts and for different conditions (i.e. chronic or acute LBP) and clinical guidelines need to be constantly updated according to literature results.

Some caution in interpreting and generalizing our results is required, particularly when interpreting change at the individual level. First, our MIC values were obtained in a convenience sample of individuals who experienced a positive outcome after a multidisciplinary rehabilitation treatment; studies are warranted in order to obtain estimates of change for deterioration of the clinical status regarding this functional parameter, as well as for other interventions, contexts, or clinical characteristics of the sample. Likewise, it is possible that different methodological and statistical approaches could lead to different results [23]. Second, some intrinsic weaknesses of GPE should be considered, e.g. the patient's ability to selectively report the effect of the intervention related to limitations in daily living activities may be influenced by additional factors (such as change in psychological factors, quality of life, etc.) [31–33].

In conclusion, our findings showed that the QBPDS is a responsive measure in Italian patients with chronic LBP undergoing multidisciplinary rehabilitation. In clinical practice and research, we recommend the use of both MICs of the QBPDS (expressing score change both in absolute value and as percentage from baseline) as a "context-specific" reference point related to minimum clinically important improvement, minding on the other hand that MDC_{95} indicates the smallest change in score that can be detected beyond random error.

References

1. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737–45.
2. de Vet HC, Terwee CB, Ostelo RW, et al. Minimal changes in health status questionnaires: Distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes. 2006;4:54.
3. Kopec JA, Esdaile JM, Abrahamowicz M, et al. The quebec back pain disability scale measurement properties. Spine (Phila Pa 1976). 1995;20(3):341–52.
4. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. Phys Ther. 2001;81(2):776–88.
5. Davidson M, Keating JL. A comparison of five low back disability questionnaires: Reliability and responsiveness. Phys Ther. 2002;82(1):8–24.
6. Mens JMA, Vleeming A, Snijders CJ, et al. Responsiveness of outcome measurements in rehabilitation of patients with posterior pelvic pain since pregnancy. Spine (Phila Pa 1976). 2002;27(10):1110–5.
7. Van Der Roer N, Ostelo RWJG, Bekkering GE, et al. Minimal clinically important change for pain intensity, functional status, and general health status in patients with nonspecific low back pain. Spine (Phila Pa 1976). 2006;31(5):578–82.
8. Demoulin C, Ostelo R, Knottnerus JA, et al. Quebec back pain disability scale was responsive and showed reasonable interpretability after a multidisciplinary treatment. J Clin Epidemiol. 2010;63(11):1249–55.

9. Wilhelm F, Fayolle-Minon I, Phaner V, et al. Sensitivity to change of the Quebec Back Pain Disability Scale and the Dallas Pain Questionnaire. Ann Phys Rehabil Med. 2010;53(1):15–23.

10. Vieira AC, Moniz S, Fernandes R, et al. Responsiveness and interpretability of the portuguese version of the quebec back pain disability scale in patients with chronic low back pain. Spine (Phila Pa 1976). 2014;39(5).

11. Speksnijder CM, Koppenaal T, Knottnerus JA, et al. Measurement properties of the Quebec back pain disability scale in patients with nonspecific low back pain: Systematic review. Phys Ther. 2016: 1816–31.

12. Monticone M, Frigau L, Mola F, et al. The Italian version of the Quebec Back Pain Disability Scale: cross-cultural adaptation, reliability and validity in patients with chronic low back pain. Eur Spine J. 2020;29(3):530–9.

13. Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: A critical review and recommendations. J Clin Epidemiol. 2000;53(5):459–68.

14. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. BMC Med Res Methodol. 2010;10:737-45.

15. Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4: European guidelines for the management of chronic nonspecific low back pain. Eur Spine J. 2006;15 (Suppl 2):S192-300.

16. Kopec JA, Esdaile JM, Abrahamowicz M, et al. The Quebec back pain disability scale: Conceptualization and development. J Clin Epidemiol. 1996;49(2):151–61.

17. Kamper SJ, Ostelo RWJG, Knol DL, et al. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol. 2010;63(7):760-6.

18. Monticone M, Ambrosini E, Rocca B, et al. A multidisciplinary rehabilitation programme improves disability, kinesiophobia and walking ability in subjects with chronic low back pain: results of a randomised controlled pilot study. Eur Spine J. 2014;23(10):2105–13.

19. Cohen J.Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.

20. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.

21. Prinsen CAC, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. Trials. 2016;17(1):449.

22. Larner AJ. Diagnostic test accuracy studies in dementia: A pragmatic approach. Diagnostic Test Accuracy Studies in Dementia: A Pragmatic Approach. 2015.

23. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008;61:102–9.

24. Schuller W, Ostelo RWJG, Janssen R, et al. The influence of study population and definition of improvement on the smallest detectable change and the minimal important change of the neck disability index. Health Qual Life Outcomes. 2014;12:53.

25. Terwee CB, Dekker FW, Wiersinga WM, et al. On assessing responsiveness of health-

related quality of life instruments: Guidelines for instrument evaluation. Qual Life Res. 2003:12:349–62.

26. Franchignoni F, Vercelli S, Giordano A, et al. Minimal clinically important difference of the disabilities of the arm, shoulder and hand outcome measure (DASH) and its shortened version (quickDASH). J Orthop Sports Phys Ther. 2014;44(1):30–9.

27. Turner D, Schünemann HJ, Griffith LE, et al. The minimal detectable change cannot reliably replace the minimal important difference. J Clin Epidemiol. 2010;63(1):28–36.

28. Revicki DA, Cella D, Hays RD, et al. Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes. 2006;4:70.

29. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: Towards international consensus regarding minimal important change. Spine (Phila Pa 1976). 2008;33(1):90–4.

30. Wang YC, Hart DL, Stratford PW, et al. Baseline dependency of minimal clinically important improvement. Phys Ther. 2011;91(5):675–88.

31. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: A review of strengths and weaknesses and considerations for design. J Man Manip Ther. 2009: 163–70.

32. Cook CE. Clinimetrics Corner: The Minimal Clinically Important Change Score (MCID): A Necessary Pretense. J Man Manip Ther. 2008;16(4):82E-83E.

33. Engel L, Beaton DE, Touma Z. Minimal Clinically Important Difference: A Review of Outcome Measure Score Interpretation. Rheum Dis Clin North Am. 2018: 177–88.

Tables

Table I. Socio-demographic and clinical characteristics of the study population (n=201).

Variable	**N.**	**%**
Sex (male/female)	80/121	39.8/60.2
Marital status		
Not married	52	25.9
Married	140	69.7
Missing	9	4.4
Occupation		
Employee	165	82.1
Self-employed	8	4.0
Housewife	27	13.4
Pensioned	1	0.5
Education		
Elementary school	9	4.5
Middle school	33	16.4
High school	96	47.8
University	61	30.3
Missing	2	1.0
Smoking		
Smoker	41	20.4
Non-smoker	160	79.6

Table II. Mean (SD) values of the QBPDS pre- vs. post-rehabilitation for the total (n=201) sample and for patients improved (GPE= 2 or 3; n=103) vs. stable (GPE= 0 or 1; n=78).

	Pre-treatment		**Post-treatment**	
QBPDS	**Mean**	**SD**	**Mean**	**SD**
Total	32.15	13.22	28.28	13.80
Improved	31.76	12.32	22.86	9.48
Stable	31.82	13.44	30.63	13.63

QBPDS: Quebec Back Pain Disability Scale; GPE: Global Perceived Effect.

Table III. Distribution-based indices of responsiveness for QBPDS.

	Cronbach's α	$ICC_{2.1}$	ES	SRM	SEM	MDC_{95}
QBPDS	0.88	0.9 (95% CI 0.86-0.93)	0.29	0.43	4.17	11.59

QBPDS: Quebec Back Pain Disability Scale; ICC: Intraclass Correlation Coefficient; ES: Effect Size; SRM: Standardized Response Mean; SEM: Standard Error of Measurement; MDC: Minimum Detectable Change.

Table IV. Mean values of the absolute and percent change from baseline, stratified by the GPE score.

	Absolute change	**Percent change**
No improvement (GPE=0) (n=40)	-1.3	-6.1%
Minimal improvement (GPE=1) (n=38)	3.8	11.8%
Moderate improvement (GPE=2) (n=54)	7.3	22.5%
Large improvement (GPE=3) (n=49)	9.6	34.0%

GPE: Global Perceived Effect

Figure Captions

Figure 1. Receiver-operating-characteristic curves of the QBPDS (n=181), showing its overall accuracy in identifying a meaningful improvement (based on absolute change score from baseline), according to the GPE at post-treatment (GPE 0 and 1 vs. GPE 2 and 3). AUC=0.83 (95% C.I. 0.75-0.90). For the optimal cut-off of 6 points:77.7% sensitivity, 80.8% specificity and 78.1% accuracy.

Figure 2. Receiver-operating-characteristic curves of the QBPDS (n=181), showing its overall accuracy in identifying a meaningful improvement (based on the percent change score from baseline), according to the GPE at post-treatment (GPE 0 and 1 vs. GPE 2 and 3). AUC=0.85 (95% C.I. 0.79-0.91). For the optimal cut-off point of 18%:80.6% sensitivity, 80.8% specificity and 79.5% accuracy.

Authors' contributions

Marco Monticone made substantial contributions to the conception and design of the work; contributed to the acquisition, analysis, and interpretation of data for the work; drafted the work and finally revised it critically for important intellectual content; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Federico Arippa made substantial contributions on the work; contributed to the analysis, and interpretation of data for the work; drafted the work and finally revised it critically for important intellectual content; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Calogero Foti made substantial contributions on the work; contributed to the interpretation of data for the work; drafted the work and finally revised it critically for important intellectual content; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Franco Franchignoni made substantial contributions to the conception and design of the work; contributed to the analysis, and interpretation of data for the work; drafted the work and finally revised it critically for important intellectual content; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

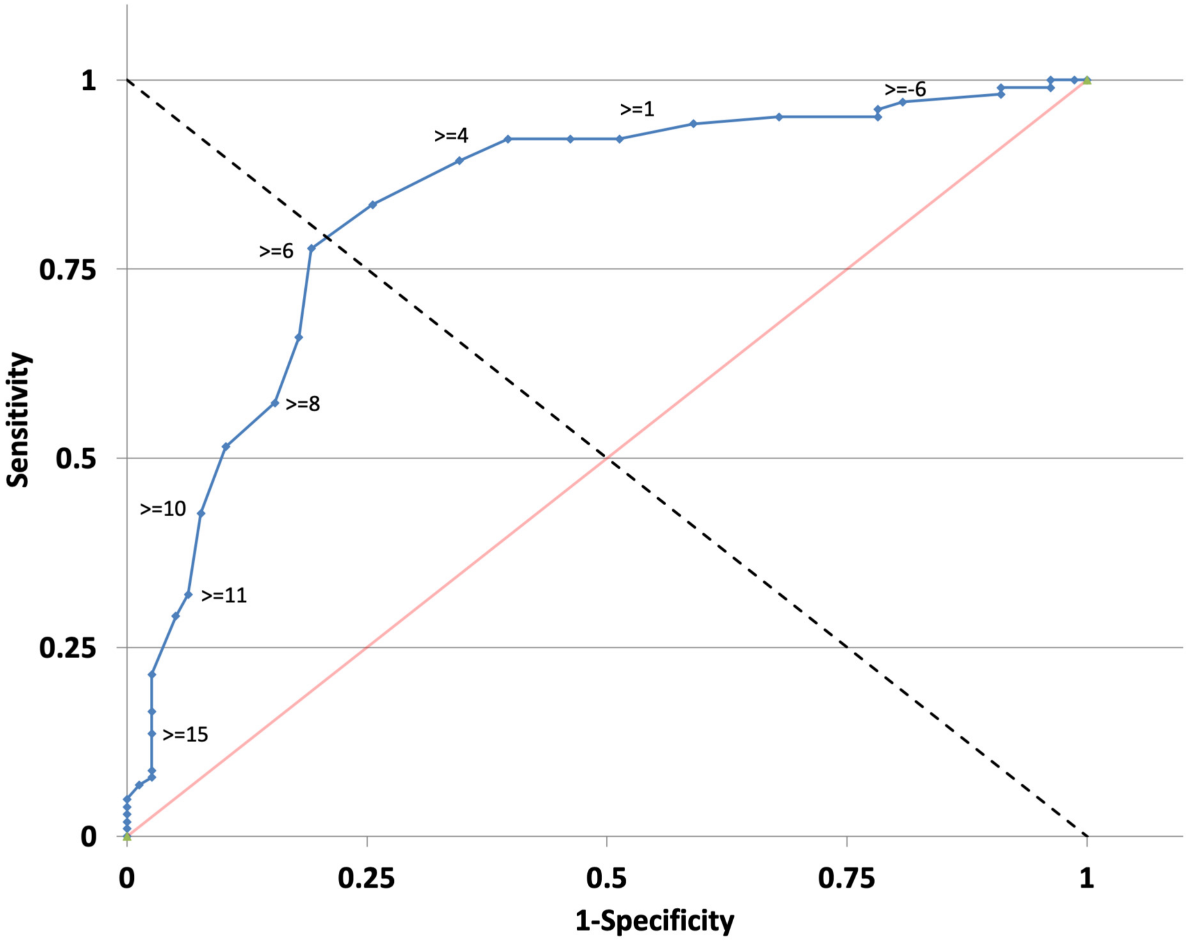
All authors read and approved the final version of the manuscript.

NOTES

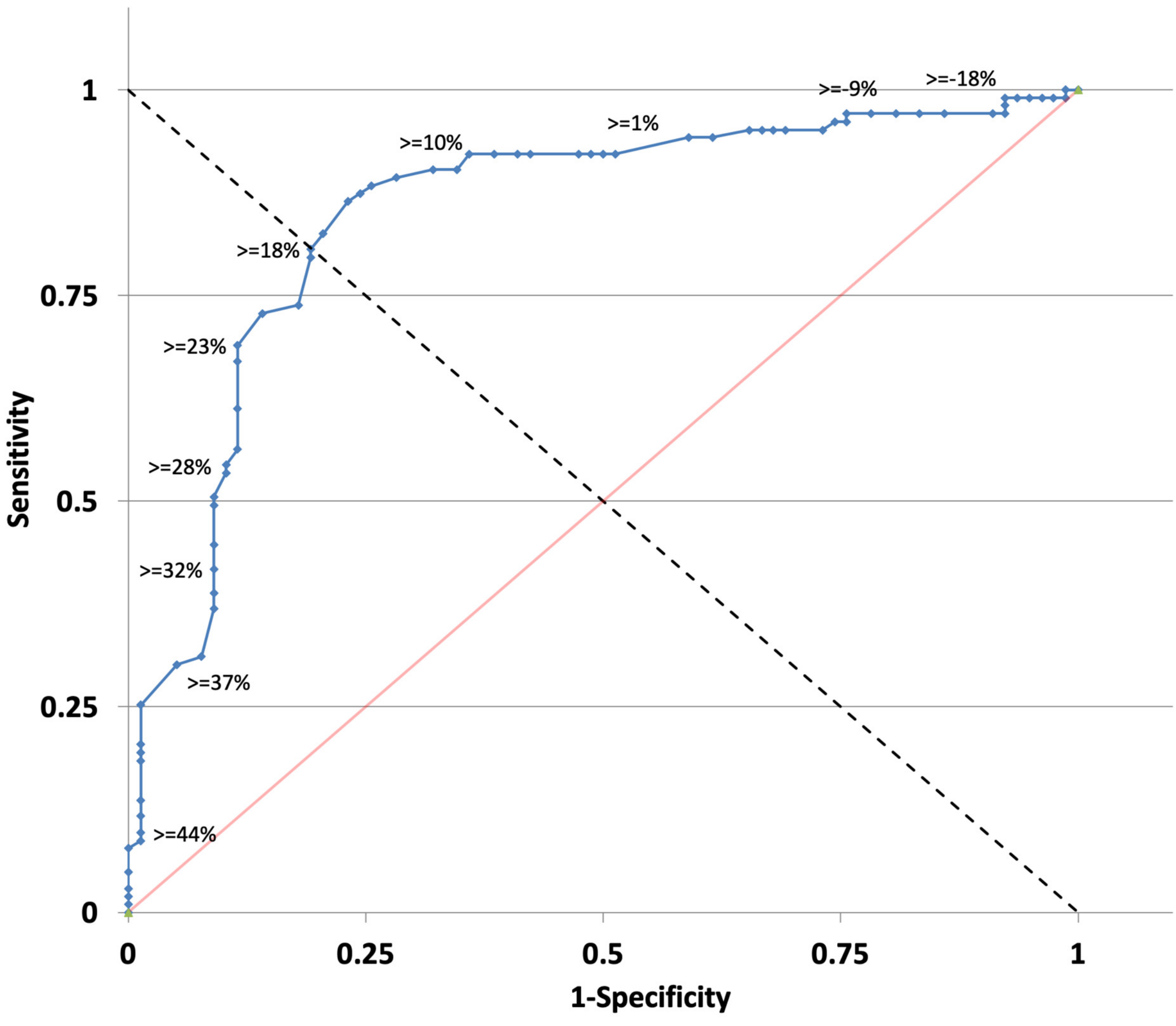
SOURCES OF GRANT SUPPORTS: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICTS OF INTEREST: The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

ACKNOWLEDGMENTS: The Authors wish to thank the hospital staff and patients who took part into the study.



1
0.75
0.5
0.25
0
Sensitivity
0
0.25
0.5
0.75
1
1-Specificity
>=-6
>=1
>=4
>=6
>=8
>=10
>=11
>=15


1
0.75
0.5
0.25
0
Sensitivity
0
0.25
0.5
0.75
1
1-Specificity
>=44%
>=37%
>=32%
>=28%
>=23%
>=18%
>=10%
>=1%
>=-9%
>=-18%