



Responsiveness of Commonly Used Patient-Reported Outcome Instruments in Lumbar Microdiscectomy

Karthik Vishwanathan¹, Ian Braithwaite²

¹Charutar Arogya Mandal, Shri Krishna Hospital, Pramukhswami Medical College, Karamsad, India

²Nuffield Hospital, Chester, UK

Study Design: Prospective observational cohort study.

Purpose: To evaluate the comparative responsiveness of Oswestry Disability Index (ODI, version 2.0), 24-item Roland-Morris Disability Questionnaire (RMDQ), 12-item Short-Form Health Survey (SF-12) physical component score (PCS), and SF-12 mental component score (MCS) in patients that underwent micro-discectomy for lumbar disc herniation.

Overview of Literature: Responsiveness is a context specific term and no study has reported the responsiveness of ODI version 2.0, RMDQ, SF-12 PCS, or SF-12 MCS in discectomy patients.

Methods: Responsiveness was assessed using distribution-based methods (effect size [ES], standardized response means [SRM], SRM difference between patients who improved and those who did not show improvement [SRM difference]), and the anchor-based method (area under the curve [AUC] of receiver operating characteristic curves). Values of ES and SRM higher than 0.8, and AUC value higher than 0.7 suggest adequate responsiveness. Outcome instrument demonstrating the highest value of SRM difference and AUC was considered the most responsive.

Results: Responsiveness was assessed in 98 participants at a mean follow-up time of 12 weeks postoperatively. The overall ES of RMDQ, ODI, SF-12 PCS, and SF-12 MCS was 2.15, 2.11, 2.08, and 0.86, respectively. The overall SRM of ODI, RMDQ, SF-12 PCS, and SF-12 MCS was 1.36, 1.43, 1.24, and 0.65, respectively. The SRM difference in RMDQ, ODI, SF-12 PCS, and SF-12 MCS was 2.64, 2.26, 1.32, and 1.29, respectively. The AUC of ODI, RMDQ, SF-12 PCS, and SF-12 MCS was 0.96, 0.96, 0.83, and 0.83, respectively.

Conclusions: ODI, RMDQ, SF-12 PCS, and SF-12 MCS demonstrated adequate responsiveness in a homogenous cohort of patients who underwent lumbar micro-discectomy. ODI and RMDQ are equally responsive and, thus, are interchangeable instruments for region specific outcomes. Both the SF-12 PCS and SF-12 MCS can assess the quality of life following lumbar micro-discectomy.

Keywords: Patient reported outcome measures; Responsiveness; Surveys and questionnaires; Outcome assessment; Sciatica; Intervertebral disc displacement; Discectomy; Quality of health care

Introduction

The Oswestry Disability Index (ODI) and the Roland-

Morris Disability Questionnaire (RMDQ) are ideal outcome instruments to assess functional outcomes; additionally, 36-item Short-Form Health Survey (SF-36) and

Received Dec 17, 2018; Revised Jan 31, 2019; Accepted Feb 14, 2019

Corresponding author: Karthik Vishwanathan

Department of Orthopaedics, Shri Krishna Hospital, Pramukhswami Medical College, Charutar Arogya Mandal, Gokalnagar, Karamsad, Gujarat, India, 388325

Tel: +91-9662088564, Fax: +91-2692-223466, +91-2692-223660, E-mail: karthikvg@charutarhealth.org

its shorter versions such as 12-item Short-Form Health Survey (SF-12) are useful instruments to assess quality of life in patients with chronic low backache due to their established validity, reliability, and responsiveness [1]. The use of SF-12 has been recommended in clinical research to evaluate generic quality of life in spine surgery patients [2]. Although the validity of an outcome instrument is an important psychometric property, reliability and responsiveness should be weighted more in the selection of an outcome instrument [1]. Responsiveness of an outcome instrument is defined as the sensitivity of the instrument to detect clinically relevant change over a period of time [3]. However, responsiveness is a context-specific term and, thus, the resulting evidence is applicable only to that particular condition and specific intervention [4]. Many studies describing the psychometric properties of commonly used patient-reported outcome instruments have considered populations with chronic nonspecific low backache. However, Glassman et al. [5] successfully argued against the use of nonspecific low backache and demonstrated differences in ODI score and SF-36 had differing baseline scores, postoperative scores, values of improvement, and pattern of improvement or deterioration when evaluating long-term follow-up for various conditions such as spondylolisthesis, scoliosis, lumbar disc pathology, instability, and lumbar spinal stenosis.

To date, no studies have been published on the comparative responsiveness of ODI, RMDQ, SF12 PCS, and SF12 MCS in patients who underwent discectomy for lumbar disc herniation. Therefore, the purpose of the study was to evaluate and report the comparative responsiveness of these instruments. Region specific outcome instruments are more responsive than generic outcome instruments [6] and, thus, we hypothesized that ODI and RMDQ would show higher responsiveness compared to SF-12 PCS and MCS. Our secondary hypothesis was SF-12 PCS would show greater responsiveness compared to SF12 MCS.

Materials and Methods

1. Data collection

This was a quantitative, single site, prospective, observational, and cohort study comparing the responsiveness of validated outcome instruments in patients who underwent lumbar microdiscectomy. Our hospital research and clinical audit department approved this study protocol.

Written informed consent was given by patients prior to surgical procedures. A prospective survey linked to service evaluation was undertaken. The study was an audit survey linked to service evaluation and approved by the research and clinical audit department of Countess of Chester Hospital, UK (approval no., 2520). As this study was an audit linked to service evaluation and improvement, institutional review board approval was waived off by the research and clinical audit department.

We recruited consecutive patients who underwent lumbar discectomy from 12th January 2009 to 4th October 2010 to the study. All the cases were either performed or supervised by one spinal surgeon. All the patients underwent non-instrumented surgical intervention (microdiscectomy) for lumbar disc prolapse and had primary complaints of leg pain which was unresponsive to non-operative methods of treatment, such as analgesics and physiotherapy. The patients had a preoperative magnetic resonance imaging (MRI) scan to confirm the diagnosis. For a patient to be considered for surgical intervention, clinical examination had to correlate with the MRI scan findings. The exclusion criteria included patients undergoing instrumented spinal surgery, decompression for tumors, trauma, and infection. Written informed consent was given by patients prior to surgical procedures.

Data related to age, sex, primary indication of surgical intervention, level of pathology and type of surgical procedure were collected on the day of the surgery. Additionally, patients completed ODI version 2.0, RMDQ, and SF-12 preoperatively in pen and paper. Six to 12 weeks after the surgery, patients returned to clinic and repeated the ODI version 2.0, RMDQ, and SF-12. Patients were asked to evaluate the change in their clinical condition and rate the success of the operative intervention.

A 6-point transition scale (external anchor) was administered to patients to evaluate clinical condition following the spinal operation. Possible options on this scale were 'cured,' 'much better,' 'a bit better,' 'the same,' 'a bit worse,' and 'much worse.' This 6-point scale was converted into binary scale for the standardized response mean (SRM) and receiver operating characteristic (ROC) analysis. Patients who responded 'cured' and 'much better' were classed as 'responders,' and the remaining answers were classed as 'non-responders.'

2. Outcome instruments

ODI version 2.0 has been shown to be valid, reliable, and responsive in chronic backache patients [7]. The ODI was calculated using the method as described previously [8]. The final score is determined as a percentage disability score, ranging from 0 (no disability) to 100 (worst disability).

The RMDQ is a valid, reliable, and responsive instrument in chronic backache [9]. For RMDQ, patients were asked to put a cross mark adjacent to the sentences that described their difficulty in daily activities due to their back pain or leg pain in the last 24 hours (today). The RMDQ is scored from 0 points (no disability) to 24 points (maximum disability).

The SF-12 is a generic instrument shown to be valid, reliable, and responsive in patients with low backache [10]. The SF-12 expresses physical component score (SF-12 PCS) and mental component score (SF-12 MCS). The value of SF-12 MCS and SF-12 PCS components range from 0 to 100 with a higher score representing better health related quality of life. The US population normalized values for SF-12 PCS and SF-12 MCS are 50.8 (standard deviation [SD]=8.9) and 50.0 (SD=9.5), respectively based on a sample of 2,105 subjects [11]. The UK norm values for SF-12 PCS and SF-12 MCS are 50.9 (SD=9.4) and 52.1 (SD=8.7), respectively based on a sample of 1,751 subjects [11]. Gandek et al. [11] concluded that there was little difference in SF-12 values of the United States and the United Kingdom, and recommended the US SF-12 version such that results from different studies can be compared. The US SF-12 (SF-12 version 1 [standard 4-week recall]) was used and PCS-12 and MCS-12 components of SF-12 were calculated. Quality Metric Health outcomes scoring software ver. 4.0 (QualityMetric Inc., Lincoln, RI, USA) was used to do this calculation.

3. Data analysis

IBM SPSS Statistics ver. 19.0 (IBM Corp., Armonk, NY, USA) was used to analyze data. The mean change in score was calculated by subtracting the mean postoperative score from the mean preoperative score. For the ODI and RMDQ, a positive value of the change score indicates improvement in clinical condition. For the SF-12, a negative value of the change score indicates improvement. Paired *t*-test was used to evaluate the statistical significance of the change in the score. The level of significance was set at 5%.

4. Distribution based methods to assess responsiveness

The effect size (ES) is the ratio between the mean change in score of an outcome instrument and the SD of the score of the outcome instrument at the baseline. The SRM is the ratio between the mean change in score of an outcome instrument and the SD of change in score of the outcome instrument. Both ES and SRM were interpreted using Cohen's criteria, where values of 0.2, 0.5, and 0.8 are considered as small, moderate, and large effects, respectively [12]. An ES and SRM value of more than 0.80 is considered optimal responsiveness of an outcome instrument [13]. SRM values between 0.2 and 0.3 are thought to suggest effects of placebo treatment, values between 0.3 and 0.5 suggesting moderately effective pharmaceutical agent, and values of more than 1.0 are usually observed after an effective surgical intervention [12].

The SRM was calculated separately for patients who responded to the surgical intervention (SRM responder) and those who did not respond to the surgical intervention (SRM non-responder). The difference between SRM responder and SRM non-responder was called 'SRM difference.' The outcome instrument that has the highest value of 'SRM difference' is considered to be the most responsive [14-16].

5. Anchor based methods to evaluate responsiveness

ROC curve was used to discriminate responders from non-responders. The area under the curve (AUC) is calculated from the ROC curve, and an AUC value of 0.5 (no discriminatory accuracy) indicates poor ability to discriminate responders from non-responders; an AUC value of 1.0 (perfect discriminatory accuracy) indicates perfect discriminatory ability of the instrument. The AUC value of <0.70, 0.70-0.79, and ≥ 0.80 are considered as indicators of suboptimal responsiveness, optimal responsiveness, and excellent responsiveness along with accurate discriminatory ability of an outcome instrument, respectively [14,15,17]. The outcome instrument having the highest AUC is considered to be the most valid and accurate discriminator of the status of change in the condition of the patient [14-16].

Table 1. Demographic and clinical features at baseline for lumbar discectomy cohort

Characteristic	Value
Total no. of cases	98
Age (yr, n=98)	44±11.1
Gender (n=98)	
Male	51 (52.0)
Female	47 (48.0)
Location of pain (n=97)	
Leg pain	17 (17.5)
Backache and leg pain	80 (82.5)
Affected extremity (n=97)	
Single leg pain	89 (91.8)
Pain in both legs	8 (8.2)
Sensory symptoms (n=93)	
Presence of paresthasias	65 (69.9)
Presence of numbness	65 (69.9)
Side of operation (n=97)	
Right	49 (50.5)
Left	45 (46.4)
Bilateral	3 (3.1)
No. of affected levels (n=98)	
Single level	96 (98.0)
Two levels	2 (2.0)
Affected levels (n=98)	
L3/L4	2 (2.2)
L4/L5	45 (45.9)
L5/S1	49 (50.0)
L4/L5+L5/S1	2 (2.0)
Type of surgery (n=98)	
Primary discectomy	85 (86.7)
Revision discectomy	13 (13.3)

Values are presented as mean±standard deviation or number of cases (%).

Results

1. Demographic characteristics

Table 1 shows the demographic and clinical features of subjects.

2. Patient response to operative intervention

The patient response rate regarding clinical condition

Table 2. Patient response and global assessment of change after operative intervention (N=93)

Outcome response	Frequency (%)
Change in clinical condition from patient perspective (external anchor)	
Patient response	
Cured	29 (31.2)
Much better	53 (57.0)
Bit better	5 (5.4)
The same	2 (2.2)
Bit worse	2 (2.2)
Much worse	2 (2.2)
Created sub-groups (for SRM difference and AUC)	
Improved (responder)	82 (88.2)
Unimproved (non-responder)	11 (11.8)

SRM, standardised response mean; AUC, area under the curve.

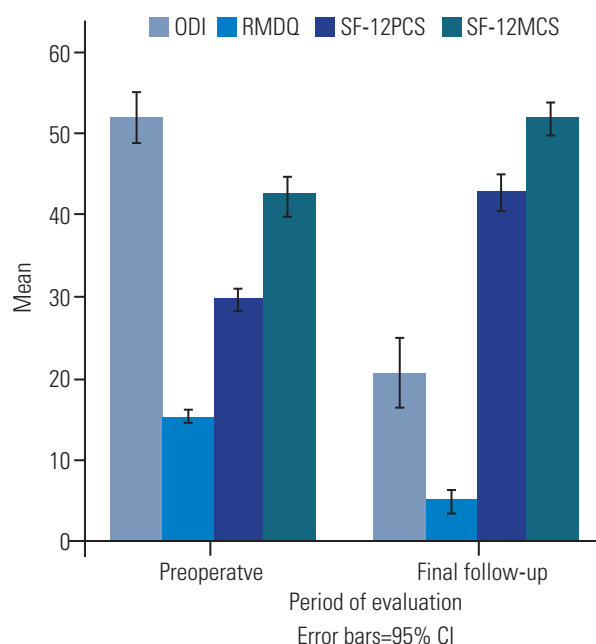


Fig. 1. Mean and 95% CIs of PRO. There was significant improvement in values of all PROs as the 95% CIs of preoperative and final follow-up scores did not overlap. CI, confidence interval; PRO, patient-reported outcome instruments; ODI, Oswestry Disability Index (version 2.0); RMDQ, 24-item Roland-Morris Disability Questionnaire; SF-12, 12-item Short-Form Health Survey; PCS, physical component score; MCS, mental component score.

subsequent to operative intervention was 93 cases (94.9%). Table 2 identifies patient response pertaining to global assessment of change after operative intervention, with 82 patients (88.2%) responding that their clinical condition

Table 3. Overall pre- and postoperative values, difference between preoperative and postoperative values, and statistical significance of the change in the value of outcome instruments

Outcome instrument	Preoperative	Postoperative	Change in score	95% Confidence interval of change in score	p-value
ODI	51.9±15.02	20.9±20.32	31.6±23.21	26.8 to 36.5	<0.0001
RMDQ	14.9±4.68	5.0±5.94	10.1±7.00	8.6 to 11.5	<0.0001
SF-12 PCS	29.5±6.41	42.7±10.33	-13.3±10.73	-15.6 to -11.0	<0.0001
SF-12 MCS	42.2±11.03	51.4±10.88	-9.5±14.56	-12.6 to -6.4	<0.0001

Values are presented as mean±standard deviation, unless otherwise stated.

ODI, Oswestry Disability Index (version 2.0); RMDQ, 24-item Roland-Morris Disability Questionnaire; SF-12, 12-item Short-Form Health Survey; PCS, physical component score; MCS, mental component score.

Table 4. SRM and AUC in different sub-groups of patient that underwent lumbar micro-discectomy

Responsiveness parameter	ODI	RMDQ	SF-12 PCS	SF-12 MCS
SRM responder	1.91	2.09	1.53	0.83
SRM non-responder	-0.35	-0.55	0.21	-0.46
SRM difference	2.26	2.64	1.32	1.29
Receiver operating curve				
AUC	0.96	0.96	0.83	0.83
95% Confidence interval of AUC	0.92 to 1.00	0.92 to 1.00	0.73 to 0.94	0.72 to 0.93
p-value	<0.0001	<0.0001	0.001	0.001

The outcome instrument having the highest value of SRM difference is considered to be the most responsive instrument. The outcome instrument having the highest value of AUC is considered to be the most responsive instrument.

SRM, standardised response mean; AUC, area under the curve; ODI, Oswestry Disability Index (version 2.0); RMDQ, 24-item Roland-Morris Disability Questionnaire; SF-12, 12-item Short-Form Health Survey; PCS, physical component score; MCS, mental component score.

improved after discectomy. Fig. 1 and Table 3 depict preoperative, postoperative, and change in values of outcome instruments, showing significant improvement all outcome instruments.

3. Distribution based method (internal responsiveness [effect size and standardized response mean])

The overall ES of ODI, RMDQ, SF-12 PCS, and SF-12 MCS was 2.11, 2.15, 2.08, and 0.86, respectively. All were large effect as per Cohen's criteria. The overall SRM of ODI, RMDQ, SF-12 PCS, and SF-12 MCS was 1.36, 1.43, 1.24, and 0.65, respectively. All were large effects as per Cohen's criteria, except for SF-12 MCS which had a moderate effect. Setting self-reported change in clinical condition as an external anchor (Table 4), the RMDQ demonstrated the highest SRM difference, and SF-12 MCS the lowest.

4. Anchor based method (external responsiveness [area under the curve for receiver operating characteristic curve])

All outcome instruments demonstrated an adequate level of external responsiveness. It was difficult to conclude if ODI or RMDQ is better as 95% confidence interval (CI) of AUC of both outcome instruments tended to overlap. Setting self-reported change in clinical condition as an external anchor (Table 4), ODI and RMDQ had near-perfect responsiveness and discriminatory validity as the upper limit of their 95% CI was 1.00 (Fig. 2). Additionally, SF-12 PCS and SF-12 MCS demonstrated identical AUC on using the change in clinical condition of the patient as the external anchor (Table 4, Fig. 2).

Discussion

Our study shows that the commonly used region-specific patient-reported functional outcome instruments and ge-

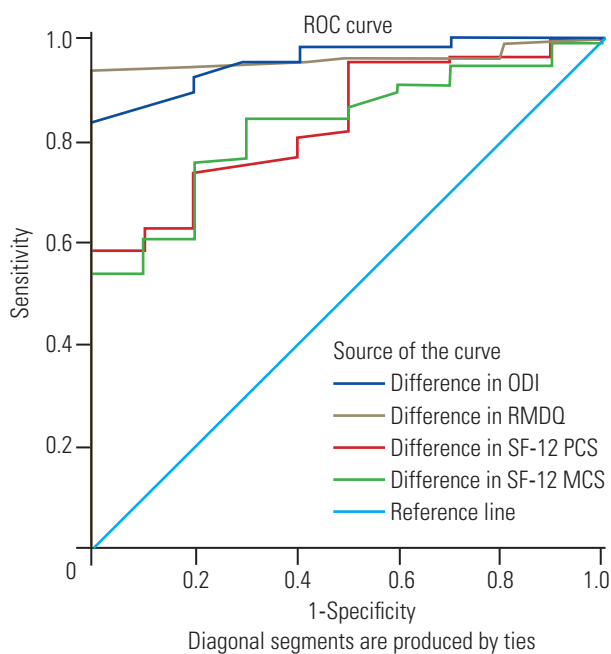


Fig. 2. Area under the curve of patient-reported outcome instruments using ROC curve using change in clinical condition from the patient's perspective as an external anchor. **ROC, receiver operating characteristic;** ODI, Oswestry Disability Index (version 2.0); RMDQ, 24-item Roland-Morris Disability Questionnaire; SF-12, 12-item Short-Form Health Survey; PCS, physical component score; MCS, mental component score.

neric quality of life outcome instruments accurately demonstrate improvement patients that underwent lumbar micro discectomy.

1. Standardized response mean difference

The SRM difference suggests that RMDQ had the highest responsiveness. As expected, RMDQ and ODI showed higher SRM difference compared to SF-12 PCS, which had a greater SRM difference than SF-12 MCS. Only one previous study has compared the SRM difference in ODI and RMDQ in a heterogenous cohort of patients with various surgically treated lumbar spine conditions such as lumbar canal stenosis, lumbar disc herniation, spondylolisthesis, and failed back syndrome [18]; this study demonstrated that the SRM difference of ODI was higher than RMDQ. The SRM difference for ODI in our study was higher than values reported previously (range, 1.11 to 2.18) [14,18]. The SRM difference for RMDQ in our study was higher than the SRM difference values quoted in literature for RMDQ (1.03) in a similar surgical cohort [18]. Only one study in a surgical cohort described an SRM difference for

SF-12 PCS and SF-12 MCS, with the SRM difference of SF-12 PCS (1.28) slightly lower than our study, but SF-12 MCS (1.44) was higher than our study [15].

2. Area under the curve

We determined that both ODI and RMDQ have near perfect discriminatory ability. SF-12 PCS and SF-12 MCS had an identical AUC and as expected, the AUC of ODI and RMDQ was higher than that of SF-12, consistent with previous studies [14,19-21], which also show higher AUC in SF-12 PCS versus SF-12 MCS. However, we show equal AUC of SF-12 PCS and SF-12 MCS, suggesting equal discriminatory capability. This is important because two previous studies showed that the AUC of SF-12 MCS was suboptimal (<0.7) [20,21], and another showed that the AUC of SF-12 PCS was suboptimal [19]. All the studies reporting the AUC of ODI [14,18-22] and AUC of RMDQ [18,23] suggested optimal responsiveness. Optimal responsiveness based on AUC of SF-12 PCS was reported in three studies [14,20,21], whereas one study reported suboptimal responsiveness [19]. Optimal responsiveness based on AUC of SF-12 MCS was reported in one study [14], whereas two studies reported suboptimal responsiveness [20,21]. Our AUC value for ODI, RMDQ, SF-12 PCS, and SF-12 MCS was comparable to published AUC values for ODI (range, 0.74 to 0.90) [14,18-22], RMDQ (range, 0.78 to 0.84) [18,23], SF12-PCS (range, 0.60 to 0.90) [14,19-21], and SF-12 MCS (range, 0.64 to 0.85) [14,20,21] in surgically treated lumbar spine conditions.

3. Effect size

The large ES ODI, RMDQ, SF-12 PCS, and SF-12 MCS suggest optimal responsiveness for all patient-reported outcome instruments in our study. Three studies showed that the ES of ODI was higher than SF-12 PCS, which was greater than SF-12 MCS, which is consistent with our findings [14,20,21]. **One study reported that ES of RMDQ suggested optimal responsiveness [18],** whereas another study reported that RMDQ had suboptimal responsiveness [23]. All the studies showed that ES of SF-12 PCS suggested optimal responsiveness [14,20,21]. One study showed that ES of SF-12 MCS showed optimal responsiveness [21], whereas two studies showed that ES of SF-12 MCS had suboptimal responsiveness [14,20]. The overall ES of ODI in our study was 2.11, which is comparable to published

reports of the ES of ODI in surgically treated lumbar spine conditions, ranging from 0.65 to 2.91 [14,18,20-22]. Three studies reported higher values of ES of ODI compared to our study [18,21,22]; **this might be explained by a difference in the duration of follow-up, different types of surgical interventions, and different types of pathology.** The overall ES of RMDQ in our study was much higher than previously reported ES of RMDQ (range, 0.76 to 0.84) after lumbar spine surgery [18,23]. One possible explanation for this discrepancy is few studies have reported the psychometric properties of RMDQ. The overall ES of SF-12 PCS and SF-12 MCS in our study was comparable to previously reported ES of SF-12 PCS (range, 0.98 to 1.70) [14,20,21] and SF-12 MCS (range, **0.63 to 0.98**) [14,20,21] in operatively treated lumbar spine conditions.

4. Standardized response mean

We showed large ESs for overall SRM values of ODI, RMDQ, and SF-12 PCS, and moderate effect for SF-12 MCS. The overall SRM of ODI was higher than the SRM of SF-12 PCS, which was higher than that of SF-12 MCS. Four similar studies [14,19-21] demonstrated that SRM of ODI was higher than SF-12 PCS and SF-12 MCS in three studies [14,20,21]. However, one study showed that SRM of ODI was lower than that of SF-12 PCS [19]. Three studies compared SRM of SF-12 PCS and SF-12 MCS [14,20,21] and surprisingly two studies reported that SRM of SF-12 PCS was lower than the SRM of SF-12 MCS [20,21]. If the value of overall SRM is greater than 0.80, it is considered as optimal responsiveness. In our study, ODI, RMDQ, and SF-12 PCS had optimal responsiveness; however, SF-12 MCS had suboptimal responsiveness. Three studies showed suboptimal ODI responsiveness [18-20] and three studies showed that SRM of ODI suggested optimal responsiveness [14,21,22]. One study reported optimal SRM of RMDQ [18], **whereas another study reported suboptimal SRM of RMDQ [23]. Three studies reported SRM values of SF-12 PCS that suggest suboptimal responsiveness [19-21] and one study reported SRM value of SF-12 PCS that suggested optimal responsiveness [14].** One study reported the SRM of SF-12 MCS to be higher than 0.80 [21], whereas two studies reported that SRM of SF-12 MCS was of suboptimal responsiveness [14,20]. The overall SRM of ODI in our study was 1.36, consistent with previously published reports of SRM of ODI (range, 0.65 to 1.90) in surgically treated lumbar spinal pathologies

[14,18,20-22]. The overall SRM of RMDQ in our work was 1.43, higher than previously published reports of SRM of RMDQ (range, 0.77 to 0.81) [18,23]. The overall SRM of SF-12 PCS and SF-12 MCS in our study was 1.24 and 0.65, respectively which was comparable to previously reported values of SRM of SF-12 PCS (range, 0.64 to 1.11) [14,19-21] and SF-12 MCS, respectively (range, 0.63 to 0.97) [14,20,21].

5. Comparison of distribution-based and anchor-based methods

Studies comparing responsiveness of ODI and RMDQ after surgical intervention are scarce. One previous study reported higher SRM difference and higher AUC of ODI suggesting that the ODI was the most responsive instrument in their study [18]. Another study showed that ODI was the most responsive instrument in their study over RMDQ in patients treated with transverse lumbar interbody fusion for degenerative spondylolisthesis [14]. The external anchors used in those studies were different from the ones used in the present study, as we use a distribution-based method (SRM difference) to determine that RMDQ was the most responsive outcome instrument. The anchor-based method (ROC) showed that both RMDQ and ODI were the most responsive instrument, suggesting that responsiveness value of an outcome instrument depends on the choice of the external anchor, clinical condition, and type of intervention. The value of responsiveness of an outcome instrument tends to vary with different external anchors [24], which might explain the discrepancy in our distribution-based and anchor-based methods.

6. Comparison of region specific and generic quality of life instruments

The responsiveness of ODI and RMDQ was higher than that of SF-12 (PCS and MCS) based on greater ES, SRM, SRM difference, and the area under the curve. This is consistent with previous observations wherein region and condition-specific outcome instruments demonstrated higher responsiveness compared to generic outcome instruments [25,26].

7. Comparison of generic quality of life instruments

In our study, responsiveness metric (SRM and ES) showed

that SF-12 PCS had optimal responsiveness as SRM, and ES were above the threshold value of 0.8; also the AUC of SF-12 was above the threshold value of 0.8 and, thus, was excellent. Additionally, we observed higher responsiveness of SF-12 PCS compared to SF-12 MCS because SF-12 PCS exhibited higher values of ES, SRM, SRM difference, and the area under the curve compared to SF-12 MCS. The SRM of SF-12 MCS was suboptimal (SRM <0.8) but the remainder of parameters met the threshold of 0.8. This was on expected line especially as SF-12 MCS is a generic quality of life instrument. Therefore, we recommend the use of SF-12 PCS and SF-12 MCS to assess the quality of life after lumbar discectomy.

8. Strengths of the study

The present study is the first study to compare responsiveness of ODI, RMDQ, and SF-12 in homogenous group of patients that underwent micro-discectomy for lumbar disc herniation. The study was methodologically robust as we used both distribution- and anchor-based methods to evaluate responsiveness with a large sample size, based on COSMIN (consensus-based standards for the selection of health measurement instruments) checklist [27].

9. Limitations of the study

Several limitations of the study warrant discussion. External validity of the present study could be affected because this was a single center study and operated or supervised by a single surgeon performing a single specific surgical procedure. Thus, these results are applicable only to lumbar disc herniation treated with lumbar discectomy, and cannot be applied to patients having backache and sciatica secondary to other causes such as spondylolisthesis or spinal canal stenosis. The study was performed as part of service evaluation audit of spinal surgery services at our district general hospital. Patients were discharged to the care of the general practitioner at either 6 weeks or 3 months depending on the clinical progress of the patient. Hence, long-term follow-up was not undertaken in the study. Finally, the assignment of subjects into responders and non-responders was arbitrary and it is possible that the significance of the difference of SRM between the two groups could be due to the simplification of division.

Conclusions

ODI, RMDQ, SF-12 PCS, and SF-12 MCS demonstrate adequate responsiveness in a homogenous cohort of patients who underwent lumbar micro-discectomy for intervertebral disc herniation. The ODI and RMDQ are equally responsive and either instrument can be used to determine region specific outcome. SF-12 PCS and SF-12 MCS can be used to assess the quality of life after lumbar micro-discectomy.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

ORCID

Karthik Vishwanathan: <https://orcid.org/0000-0001-7099-941X>

Author Contributions

Conception and design: KV, IB; data collection and assembly of data: KV, IB; provision of patients: IB; administrative support: IB; data analysis and interpretation: KV, IB; manuscript drafting: KV; and critical revision of manuscript and supervision: IB

References

1. Chapman JR, Norvell DC, Hermsmeyer JT, et al. Evaluating common outcomes for measuring treatment success for chronic low back pain. *Spine (Phila Pa 1976)* 2011;36:S54-68.
2. DeVine J, Norvell DC, Ecker E, et al. Evaluating the correlation and responsiveness of patient-reported pain with function and quality-of-life outcomes after spine surgery. *Spine (Phila Pa 1976)* 2011;36:S69-74.
3. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
4. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102-9.

5. Glassman SD, Carreon LY, Djurasovic M, et al. Lumbar fusion outcomes stratified by specific diagnostic indication. *Spine J* 2009;9:13-21.
6. Suk M, Norvell DC, Hanson B, Dettori JR, Helfet D. Evidence-based orthopaedic surgery: what is evidence without the outcomes? *J Am Acad Orthop Surg* 2008;16:123-9.
7. Leclaire R, Blier F, Fortin L, Proulx R. A cross-sectional study comparing the Oswestry and Roland-Morris Functional Disability scales in two populations of patients with low back pain of different levels of severity. *Spine (Phila Pa 1976)* 1997;22:68-71.
8. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)* 2000;25:2940-52.
9. Stratford PW, Binkley JM. Applying the results of self-report measures to individual patients: an example using the Roland-Morris Questionnaire. *J Orthop Sports Phys Ther* 1999;29:232-9.
10. Luo X, George ML, Kakouras I, et al. Reliability, validity, and responsiveness of the short form 12-item survey (SF-12) in patients with back pain. *Spine (Phila Pa 1976)* 2003;28:1739-45.
11. Gandek B, Ware JE, Aaronson NK, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project: International Quality of Life Assessment. *J Clin Epidemiol* 1998;51:1171-8.
12. Schouffoer AA, van der Giesen FJ, Beart-van de Voorde LJ, Wolterbeek R, Huizinga TW, Vliet Vlieland TP. Validity and responsiveness of the Michigan Hand Questionnaire in patients with systemic sclerosis. *Rheumatology (Oxford)* 2016;55:1386-93.
13. Naal FD, Impellizzeri FM, Rippstein PF. Which are the most frequently used outcome instruments in studies on total ankle arthroplasty? *Clin Orthop Relat Res* 2010;468:815-26.
14. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, Glassman SD, McGirt MJ. Accurately measuring the quality and effectiveness of lumbar surgery in registry efforts: determining the most valid and responsive instruments. *Spine J* 2014;14:2885-91.
15. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, McGirt MJ. Accurately measuring the quality and effectiveness of cervical spine surgery in registry efforts: determining the most valid and responsive instruments. *Spine J* 2015;15:1203-9.
16. Zuckerman SL, Chotai S, Devin CJ, et al. Surgical resection of intradural extramedullary spinal tumors: patient reported outcomes and minimum clinically important difference. *Spine (Phila Pa 1976)* 2016;41:1925-32.
17. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007;7:541-6.
18. Mannion AF, Junge A, Grob D, Dvorak J, Fairbank JC. Development of a German version of the Oswestry Disability Index: part 2: sensitivity to change after spinal surgery. *Eur Spine J* 2006;15:66-73.
19. Parker SL, Adogwa O, Mendenhall SK, et al. Determination of minimum clinically important difference (MCID) in pain, disability, and quality of life after revision fusion for symptomatic pseudoarthrosis. *Spine J* 2012;12:1122-8.
20. Parker SL, Mendenhall SK, Shau D, et al. Determination of minimum clinically important difference in pain, disability, and quality of life after extension of fusion for adjacent-segment disease. *J Neurosurg Spine* 2012;16:61-7.
21. Parker SL, Mendenhall SK, Shau DN, et al. Minimum clinically important difference in pain, disability, and quality of life after neural decompression and fusion for same-level recurrent lumbar stenosis: understanding clinical versus statistical significance. *J Neurosurg Spine* 2012;16:471-8.
22. Parker SL, Adogwa O, Paul AR, et al. Utility of minimum clinically important difference in assessing pain, disability, and health state after transforaminal lumbar interbody fusion for degenerative lumbar spondylolisthesis. *J Neurosurg Spine* 2011;14:598-604.
23. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995;48:1369-78.
24. Lauridsen HH, Hartvigsen J, Korsholm L, Grunnet-Nilsson N, Manniche C. Choice of external criteria in back pain research: does it matter?: recommendations based on analysis of responsiveness. *Pain* 2007;131:112-20.
25. Wright JG, Young NL. A comparison of different indices of responsiveness. *J Clin Epidemiol* 1997;50:239-46.
26. Wiebe S, Guyatt G, Weaver B, Matijevic S, Sidwell

- C. Comparative responsiveness of generic and specific quality-of-life instruments. *J Clin Epidemiol* 2003;56:52-60.
27. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651-7.