Journal of Clinical and Translational Research 2022; 8(6): 594-601

Journal of Clinical and Translational Research

Journal homepage: http://www.jctres.com/en/home

ORIGINAL ARTICLE

Accuracy of post-operative recall by degenerative cervical myelopathy patients using the modified Japanese Orthopaedic Association scale

Shuai Chang[1,2,3†], Nanfang Xu[1,2,3†], Yubo Luo[4], Shaobo Wang[1,2,3*], Zhongjun Liu[1,2,3*]

[1]Department of Orthopedics, Peking University Third Hospital, Beijing, 100191 China, [2]Beijing Key Laboratory of Spinal Disease Research, Peking University Third Hospital, Beijing, 100191 China, [3]Engineering Research Center of Bone and Joint Precision Medicine, Ministry of Education, Peking University Third Hospital, Beijing, 100191 China, [4]School of Mathematics and Statistics, Beijing Technology and Business University, Beijing, 102488 China

[†]These authors contributed equally to this work.

ARTICLE INFO

Article history:
Received: October 7, 2022
Revised: November 1, 2022
Accepted: November 3, 2022
Published online: November 15, 2022

Keywords:
anterior cervical discectomy
degenerative cervical myelopathy
intervertebral fusion
neurological function
patient-reported outcomes
recall error

**Corresponding authors:*
Shaobo Wang, Department of Orthopedics, Peking University Third Hospital, Beijing, 100191 China, Beijing Key Laboratory of Spinal Disease Research, Peking University Third Hospital, Beijing, 100191 China Engineering Research Center of Bone and Joint Precision Medicine, Ministry of Education, Peking University Third Hospital, Beijing, 100191 China.
Email: bysywangshaobo@outlook.com
Zhongjun Liu, Department of Orthopedics, Peking University Third Hospital, Beijing, 100191 China, Beijing Key Laboratory of Spinal Disease Research, Peking University Third Hospital, Beijing, 100191 China, Engineering Research Center of Bone and Joint Precision Medicine, Ministry of Education, Peking University Third Hospital, Beijing, 100191 China.
Email: zjliu@bjmu.edu.cn

ABSTRACT

Background and Aim: The modified Japanese Orthopaedic Association (mJOA) scale is one of the primary measures of neurological function used on patients with degenerative cervical myelopathy (DCM). Contrary to some reports, the mJOA is not based on patient-reported outcomes as it is an assessment conducted by physicians, allied health professionals, or trained staff. To date, the accuracy of post-operative recall by DCM patients of their pre-operative neurological function, as assessed by the mJOA scale, has not been examined. This study, therefore, aimed to evaluate recall accuracy in DCM patients using the mJOA scale.

Methods: This study analyzed recall capacity of DCM patients who had undergone anterior cervical discectomy and fusion by a single surgeon at a large academic spine center between February 2012 and August 2017. Patient recall of neurological function pre-surgery was assessed at 3, 12, and 24 months post-surgery using the mJOA scale. Actual mJOA scores were also determined at each follow-up. Recall error (RE) was defined as the difference between recalled mJOA score at each post-operative visit and the actual baseline score. Age, gender, surgical segments, hospital length of stay, actual mJOA scores at follow-up, and actual rate of improvement in mJOA score were analyzed as predictors of recall accuracy. Descriptive statistics were collected to profile the characteristics of patients enrolled in the study cohort. All statistical computing and graphing were performed with R software and generalized estimating equation (GEE) model fitting was done using geepack package.

Results: A total of 105 patients (56.2% of males and 43.8% of females) were enrolled in the study. The median ± SD (range) age at the pre-surgical baseline measurement was 50 ± 8 (25 – 78) years. The recalled mJOA scores at the three follow-up time points were lower than the actual mJOA scores. The recall accuracy gradually decreased over time. Estimated coefficients showed that all variables in the GEE model except for surgical fusion segments were significant ($P < 0.05$). The pre-operative actual baseline mJOA score was inversely associated with RE. An increasing actual mJOA score over time had a significant positive influence on RE. Greater RE was found in males compared to females. Unexpectedly, age was inversely associated with RE.

Conclusions: The RE increases with the time interval between pre-surgical measurement and post-surgical follow-up and is more prominent in male DCMs patients following upper spine surgery.

Relevance for Patients: It is necessary to select post-operative patients who need to pay attention according to the three factors of post-operative time, gender, and age, that is, patients with large RE should be given early or timely psychological counseling and treatment concerns, so as to reduce the occurrence of potential medical disputes and improve the level of medical safety.

DOI: http://dx.doi.org/10.18053/jctres.08.202206.012

1. Introduction

Patient-reported outcomes (PROs) play a crucial role in quantifying post-operative pain and assessing a patient's quality of life after surgery [1]. However, while PROs facilitate and improve quantitative data collection, their inherent subjectivity exposes PROs to individual interpretation, inaccurate self-reporting, and the patient's potential inability to recall pre-intervention injury [2].

The modified Japanese Orthopaedic Association (mJOA) score is an investigator-administered tool used to evaluate neurological function in patients with degenerative cervical myelopathy (DCM). It is an 18-point scale that addresses upper (5 points) and lower extremity (7 points) motor function, sensation (3 points), and micturition (3 points). In the literature, there are reports of a patient-derived version of the mJOA for the patients to report their conditions. There are also publications on validating the PRO version of the JOA score. Although the PRO-JOA was deemed comparable to the mJOA, the PRO-JOA and mJOA scores should be regarded as different outcomes. The mJOA score is not a PRO in that it is an assessment conducted by physicians, allied health professionals, or trained staff. Furthermore, there is controversy about whether the mJOA should be administered retrospectively to patients in the absence of a baseline score to ascertain their recollection of the pre-operative state.

Recall bias is a well-known source of systematic error in clinical research. It is a differential form of misclassification bias and the risk estimation may deviate or move to zero [2,3]. Although many orthopedic studies have reported recall bias [4-9], only few studies have reported the accuracy of recall of PROs in patients with DCM. The purpose of this study was therefore to (1) study the accuracy of patients' recall of pre-operative neurological function after anterior cervical discectomy and fusion (ACDF); (2) describe the direction and magnitude of bias; and (3) evaluate the overall consistency between the actual mJOA score and the recall mJOA score in a Chinese patient cohort. We used the mJOA scale to perform an accuracy study of pre-operative neurological function recall in patients with DCM who had undergone ACDF.

2. Methods

2.1. Study design

This study analyzed retrospective data on DCM patients at a large academic spine center where ACDF was performed by a single surgeon between February 2012 and August 2017. Patient recall of neurological function at 3, 12, and 24 months after surgery was compared to pre-operative actual baseline scores according to the mJOA scale (1994 version). Our mJOA comprises a validated 17-point measure of neurological function in this patient population, with 0 indicating the most significant functional impairment and 17 indicating no neurological debilitation (see online data supplement).

Recall error (RE) was defined as the difference between recalled mJOA score and the actual baseline score at each post-operative visit (POV) and analyzed for age, gender, surgical segments, length of stay, current actual mJOA scores, and actual rates of improvement in patient recall accuracy. Approval from the Peking University Third Hospital medical ethics committee was obtained before initializing the study under approval number M2020380.

2.2. Inclusion and exclusion criteria

The inclusion and exclusion criteria are presented in Table 1.

2.3. Data collection

The four key time points in the study were 3 days before surgery and 3, 12, and 24 months after surgery. The mJOA score was calculated at the four key time points after the patients had filled out the questionnaire. A baseline mJOA score was obtained within 3 days before surgery as part of standard care. At 3, 12, and 24 months after surgery, patients were asked to recall their baseline mJOA score. Furthermore, the actual mJOA scores reflecting the patient's functional status at the time of measurement were recorded. Descriptive statistics were collected to describe the characteristics of patients enrolled in the study cohort.

The two-sided paired t-test and the Wilcoxon signed-rank test were used to evaluate the mean difference between the recalled and pre-operative actual baseline mJOA. The Pearson correlation coefficient was calculated to determine the correlation between recalled and pre-operative actual baselines JOA, with 0 representing no correlation and 1 representing perfect correlation.

RE, calculated as the difference between recalled and pre-operative actual baseline mJOA (i.e., recalled mJOA minus pre-operative actual baseline mJOA), was used to measure recall accuracy. Positive RE indicates that the numerical value of the recalled mJOA minus the pre-operative actual baseline mJOA was greater than the actual score, while negative RE indicates the opposite. For those who recalled a baseline mJOA equal to the actual score, the RE was zero.

Generalized estimating equation (GEE) was developed by Liang and Zeger (1986) [10] to produce more efficient and unbiased regression estimates when analyzing longitudinal or repeated measures research designs with non-normal response variables. GEE can cope with correlated data within subjects [10]. The greatest advantage of the GEE model is that there is no need to specify the whole distribution of the response. Only the mean

Table 1. Inclusion and exclusion criteria of patients with DCM after ACDF surgery

Inclusion criteria	Exclusion criteria
• Diagnosis of cervical spondylotic myelopathy	• Only radiculopathy symptoms, symptoms or deterioration due to cervical trauma
• Clinical symptoms of cervical spondylotic myelopathy	• Radiological examination of the posterior longitudinal ligament, with neoplastic, infectious, or metabolic disease
• Pre-operative mJOA score of ≤ 13	• History of cervical surgery or cervical deformity
• Radiological examination with degenerative cervical stenosis and cervical cord decompression	• Underwent surgery other than ACDF (e.g., simultaneous vitrectomy, laminoplasty, etc.)
• Aged ≥ 18 years at the time of surgery	• Loss or inability to be interviewed during follow-up, incomplete documentation

DCM: Degenerative cervical myelopathy, ACDF: Anterior cervical discectomy and fusion, mJOA: modified Japanese Orthopaedic Association

DOI: http://dx.doi.org/10.18053/jctres.08.202206.012

structure, the mean-variance relationship, and specification of the covariance structure need to be defined [11].

We used GEE [12] to detect the association between a series of predictors and RE. In our model, the dependent variable was the mJOA RE. Independent variables included (1) the pre-operative actual baseline mJOA, (2) the actual mJOA on each visit, (3) gender, (4) age at the time of surgery, and (5) time between baseline and follow-up (in months). The GEE was as follows:

$$RE(RecallError_{it}) = \beta_0 + \beta_1 JOAPreop_{it} + \beta_2 JOAActual_{it} + \beta_3 SurgeryNo_{it} + \beta_4 Visit_{it} + \beta_5 Age_{it} + \beta_6 Gender_{it}; \quad t = 3, 12, 24; \; i = 1,\ldots,n$$

Where JOAPreop represents pre-operative baseline mJOA score, JOAActual represents actual mJOA score at each visit, SurgeryNo represents surgical fusion segments, age is age at time of surgery, and visit represents the time between baseline measurement and follow-up (in months). Because we found no evidence indicating notable interaction between the independent variables, interaction terms were not included in our model. For the working correlation structure in GEE, we considered both exchangeable working correlation and AR(1) working correlation. All statistical computing and graphing were done in the R platform [13]. GEE model fitting and diagnostics were performed with the glmtoolbox package [14].

3. Results

A total of 105 patients (56.2% of males and 43.8% of females) were enrolled in the study from February 2012 to August 2017. The median ± SD (range) age at baseline was 50 ± 8 (25 – 78) years. Figure 1 displays the boxplots of recalled and actual mJOA scores at the 3 visits (at 3, 12, and 24 months follow-up). For boxplots, the recalled mJOA score at three follow-up time points was less than the actual mJOA. The dispersion of recalled mJOA scores became larger over time, meaning that recall accuracy gradually decreased.

Figures 2-4 show scatter plots of recalled versus actual mJOA scores at each visit. Figure 5 presents the RE for the three visits.

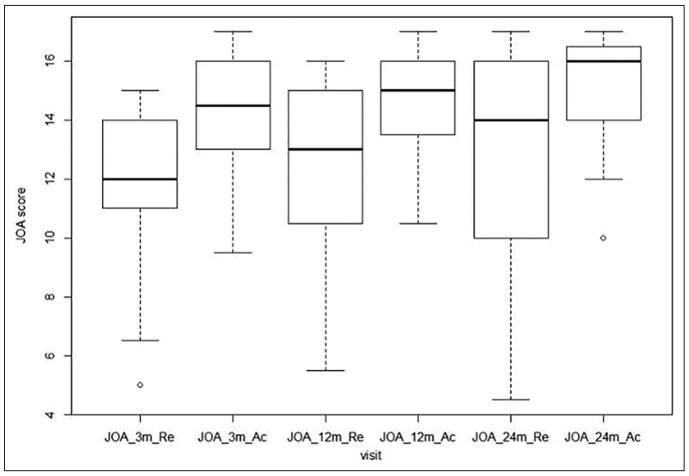

Figure 1. Recalled and actual modified Japanese Orthopaedic Association (mJOA) scores at the three visits. JOA_3m_Re, JOA_12m_Re, and JOA_24m_Re signify recalled mJOA scores at 3, 12, and 24 months, respectively. JOA_3m_Ac, JOA_12m_Ac, and JOA_24m_Ac signify actual mJOA scores at 3, 12, and 24 months, respectively.

Figure 6 represents a scatterplot of RE versus actual mJOA score change by gender. Table 2 provides results of correlation test between recalled and actual mJOA scores, which were significantly correlated at each visit. The Pearson correlation coefficients between recalled errors for three visits are given in Table 3, indicating that there was a time effect on REs and that the

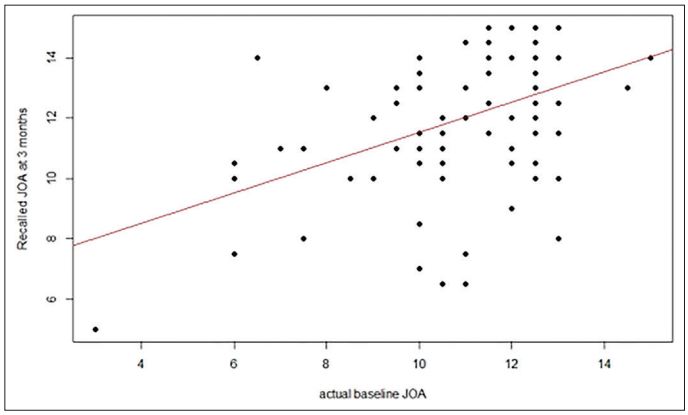

Figure 2. Scatter plot of recalled modified Japanese Orthopaedic Association (mJOA) scores at 3 months versus actual mJOA scores.

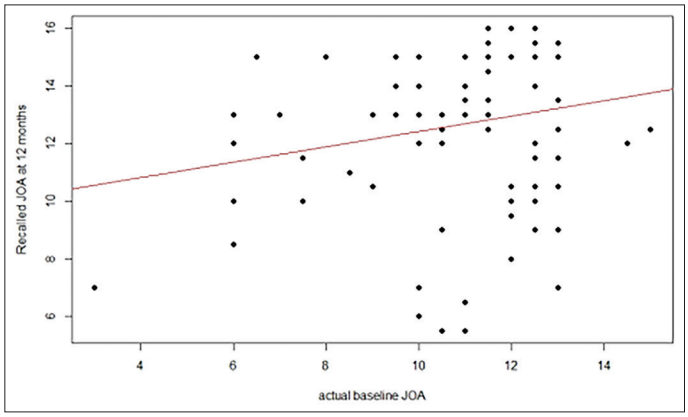

Figure 3. Scatter plot of recalled modified Japanese Orthopaedic Association (mJOA) scores at 12 months versus actual mJOA scores.

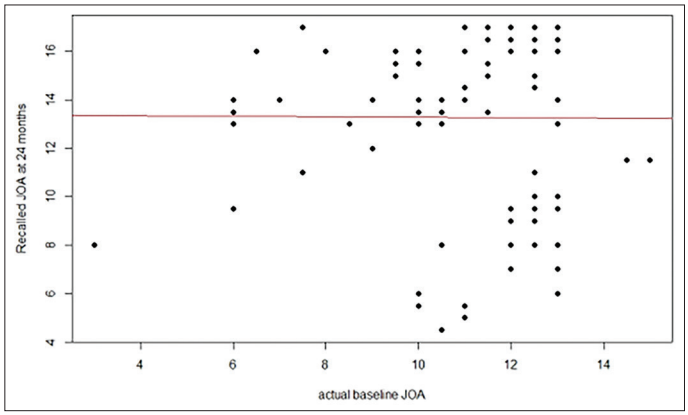

Figure 4. Scatter plot of recalled modified Japanese Orthopaedic Association (mJOA) scores at 24 months versus actual mJOA scores.

DOI: http://dx.doi.org/10.18053/jctres.08.202206.012

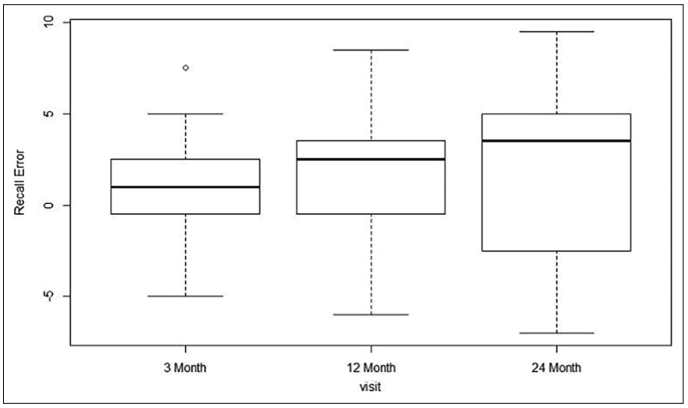

Figure 5. Recall error of the pre-operative actual baseline modified Japanese Orthopaedic Association score at each visit.

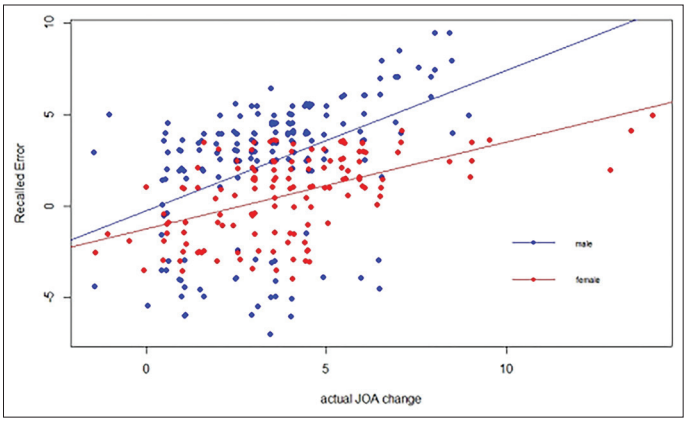

Figure 6. Scatter plot of recall error versus actual modified Japanese Orthopaedic Association score change stratified by gender.

AR(1) working correlation was probably more preferred. Table 4 presents the result analysis of the GEE model.

In the first step of our model fitting procedure, we fit the GEE model for two working correlation matrices, namely, exchangeable and AR(1). We found that there were no significant differences between these two settings; the QIC was 2177.62 and 2181.72, respectively, and the coefficients were similar. The estimated correlation parameters were 0.8499 with a standard error of 0.0266 for AR(1) working correlation and 0.8106 with a standard error of 0.0324 for exchangeable working correlation. On inspection of the residuals yielded by those models, we found several outliers.

In the second step, we excluded the outliers with residuals <−6 retrieved in the first step. A total of 96 subjects were retained and nine subjects were excluded. Subsequently, the model with AR(1) working correlation was fitted. The refitted model's QIC was now 854.26, which had a significant decrease from the first step. The estimated correlation parameter was 0.6767 with a standard error of 0.0660.

Table 4 provides the results for the refitted model. As becomes evident from Table 4, all variables except SurgeryNo were significant ($P < 0.05$). The most significant variables were JOAPreop, visit, gender, and age. All estimated coefficient signs were consistent with expectations. The greater the pre-operative actual baseline mJOA score, the smaller the RE. Actual mJOA score increases had a significant positive influence on RE, but its effects were less than the baseline mJOA score. The RE increased with follow-up interval length, as was expected. Greater RE was found in males than in females. The variable age had a negative effect on RE, indicating that elderly patients exhibited smaller RE.

Table 2. Pearson correlation between recalled and actual baseline mJOA score.

Follow-up time	Pearson's *r*
3 months	0.4545 ($P = 0.0000$)
12 months	0.1964 ($P = 0.0445$)
24 months	−0.0048 ($P = 0.9612$)

mJOA: modified Japanese Orthopaedic Association

Table 3. Pearson correlation between recall errors of three visits.

Comparison groups	Pearson's *r*
3 months versus 12 months	0.9498 ($P = 0.000$)
3 months versus 24 months	0.9177 ($P = 0.000$)
12 months versus 24 months	0.9679 ($P = 0.000$)

Table 4. Result analysis of the GEE model.

Coefficient	Estimate	Standard error	Wald	*P* (>|W|)
Intercept	12.0341	1.7914	45.1261	0.0000
mJOAPreop	−0.6670	0.0867	59.1407	0.0000
mJOAActual	0.1785	0.0607	8.6548	0.0033
SurgeryNo	0.3354	0.2292	2.1429	0.1432
Visit	0.0631	0.0108	34.3728	0.0000
Age	−0.0518	0.0238	4.7530	0.0292
Gender	−2.6695	0.3350	63.4864	0.0000

mJOA: modified Japanese Orthopaedic Association, GEE: Generalized estimating equation

4. Discussion

Patients with DCM recalled pre-operative neurological function during 2 years following ACDF surgery using the mJOA scoring system. From 3 to 24 months, patients' recall of pre-operative neurological function grew increasingly erroneous, showing a definite time effect. The recall deviation was smaller at 3 months than at 12 months and 24 months. The findings of this study indicated that, on the one hand, the accuracy of recall differed abated with follow-up time, which echoes several other studies [4,7,15]. On the other hand, relying on patient recall did not provide an accurate assessment of the pre-operative status, which is also consistent with several other studies assessing patient recall bias [5,6,8,9,16-18]. Some scholars also pointed out that recall accuracy may be determined by patient characteristics as well as characteristics of the exposure of interest [3,8,19]. In a cohort of patients diagnosed with anterior cruciate ligament injury and who were followed for 2 years, Randsborg *et al.* [20] noted that the recall bias of pre-injury knee function following anterior cruciate ligament reconstruction was small and not clinically meaningful for the majority of patients. However, patients with

DOI: http://dx.doi.org/10.18053/jctres.08.202206.012

poor outcome had a clinically relevant and significant recall bias. Coincidentally, in a cohort of back pain patients followed for 10 years, Dawson *et al.* [8] found that radicular symptoms, frequency and location of pain, and activities affecting pain were recalled with greater accuracy than specific qualities of pain such as severity. Our study focused on the total score of the mJOA and did not refine the score of each subgroup, so no more detailed conclusions could be drawn. However, the aforementioned studies revealed that by refining the grouping of the study, broadening of patient characteristics, and fully exposing the research subjects, the recall bias would show different characteristics. Pellisé *et al.* reported significantly worse recalled pre-operative symptoms in patients undergoing spinal fusion for lower back pain [9]. Lowe *et al.* [18] showed that patients were able to recall pre-operative function with considerable accuracy for up to 12 months after total shoulder arthroplasty. However, beyond 6 weeks postoperatively, patients recalled having worse pain than they originally reported. This trend was also confirmed in the present study. Boxplots of the study in Figure 1 showed that the recalled mJOA scores at three follow-up time points were less than the actual mJOA scores and the dispersion of recalled mJOA scores became vaster over time, meaning that recall accuracy gradually decreased.

The larger the pre-operative baseline mJOA, the smaller the RE. Actual mJOA score increases positively affected RE, although less than the baseline mJOA. The baseline value affected accuracy. This may be explained by the possibility that individuals with a high pre-operative baseline value had less room for post-operative mJOA score improvement, and memory error was small.

Males had worse recall than females. On the one hand, in a multicenter MRI study of mild traumatic brain injury, Shetty *et al.* [21] pointed out that the female gender was associated with an increase in symptom severity scores at every time point when studying RE. In other words, women had a greater RE than men. Although the above study [21] also focused on recall bias, the conclusions based on gender are contrary to the present study, and the reason for the analysis may be that the different characteristics of the study subjects caused the differences in findings. On the other hand, Holden *et al.* [22] pointed out that women placed relatively more emphasis on categorical cues, while men relied more heavily on metric information when they examined ambidextrous individual differences in the relative weighting of these cues in spatial location memory. This study showed why recall in men was better than in women. Men depended more on references and relative evidence for recall, whereas women looked for direct or absolute evidence, resulting in less recall bias in women.

Three limitations should be noted in the present study. First, although our pre-operative and post-operative mJOA scores were completed face to face and further evaluated in the investigator-guided patient setting, they were prone to responder bias because of the retrospective nature of the study [23]. The previous psychometric studies showed that retrospective studies were not more accurate than the differences in prospective recordings [3,24]. In a retrospective study, the treatment responders had the advantages of recalibrated post-operative self-reported results compared with non-responders [18]. Meanwhile, this shift in response might evolve over time because the personal experiences of the patient, where their own internal standards and expectations of treatment effects may play an important role [25]. This was one of the systematic biases, but it might be eliminated in prospective studies. To overcome these biases, recall adjustment calculation and sensitivity analysis were recommended in some studies [26]. These highlighted the importance of collecting PROs prospectively and not retrospectively. Further prospective studies are needed to evaluate the effectiveness of these change score calculations for mJOA scores.

Second, relevant literature [27] revealed that the Oswestry Disability Index (ODI) score has been proven to serve more like a function of content than the format presented by the project. When patients were asked to fill in the ODI questionnaire repeatedly at different follow-up time points, they might have "learned" the ODI-related problems, resulting in systematic bias. However, during treatment, the ODI's repeated disability assessment had stable psychometric characteristics, so it would not be the main source of bias [27]. Since there was no part on psychological measurement in the mJOA score, patients who had "learned" in the previous follow-up might have been primed for the abovementioned systematic bias. The future direction is to add the appropriate psychological measurement questionnaire to the research to avoid systematic bias to the greatest possible extent.

Third, our study focused specifically on longer-term follow-up (at least 2 years postoperatively), with the first follow-up time point being 3 months after surgery. However, some studies have pointed out that the accuracy of recall was higher when the post-operative follow-up time was 1 – 7 days [28,29]. In addition, Marsh *et al.* [15] pointed out that patients undergoing total hip arthroplasty could accurately recall their pre-operative health status at 6 weeks postoperatively. The factor "Age" had negative effects on RE, meaning that older patients exhibited smaller RE. There were no relevant conclusions in published literature for us to compare. Accordingly, it is possible that DCM patients might have greater recall accuracy at shorter-term follow-up (e.g., up to 6 weeks), which was not probed in this study. Future recall bias studies should therefore include shorter time frames.

5. Conclusions

There was a significant time effect on the recall accuracy of mJOA scores at follow-up in that the RE increased with follow-up time interval. Greater RE was found in males compared to females.

The research on the accuracy of RE was conducted mainly to understand the retrospective ability and accuracy of DCM patients' pre-operative neurological function after surgery. We found that the post-operative recall time and male population were negative predictors of recall accuracy.

During post-operative follow-up, we met with patients who suffered from unsatisfactory neurological function recovery and no obvious symptom relief. These visits required us to invest much patience and energy to help the patients recall their pre-operative neurological function as much as possible. By browsing the data of

DOI: http://dx.doi.org/10.18053/jctres.08.202206.012

the previous mJOA scores, the accurate pre-operative neurological status of the patient could be retrieved. By comparing the data to the post-operative results, the patient can more intuitively and objectively understand the real changes as a result of the surgery. This could help patients to become more appreciative of the surgical intervention. Due to this study, we recommend that the abovementioned patient groups require more healthcare-related attention to reduce the occurrence of potential medical disputes and improve medical safety by placating patients' dissatisfaction or anxiety.

In future studies, we need to include patients with posterior cervical spine surgery and even lumbar spine surgery, expand the study scope, and add different types of surgery as a factor. Future studies should preferably be carried out prospectively to reduce selection bias.

Acknowledgments

None.

Funding

None.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] McCormick JD, Werner BC, Shimer AL. Patient-Reported Outcome Measures in Spine Surgery. J Am Acad Orthop Surg 2013;21:99-107.

[2] Stepan JG, London DA, Boyer MI, Calfee RP. Accuracy of Patient Recall of Hand and Elbow Disability on the QuickDASH Questionnaire Over a Two-Year Period. J Bone Joint Surg Am 2013;95:e176.

[3] Coughlin SS. Recall Bias in Epidemiologic Studies. J Clin Epidemiol 1990;43:87-91.

[4] Howell J, Xu M, Duncan CP, Masri BA, Garbuz DS. A Comparison between Patient Recall and Concurrent Measurement of Preoperative Quality of Life Outcome in Toal Hip Arthroplasty. J Arthroplasty 2008;23:843-9.

[5] Lingard EA, Wright EA, Sledge CB, Kinemax Outcomes Group. Pitfalls of using Patient Recall to Derive Preoperative Status in Outcome Studies of Total Knee Arthroplasty. J Bone Joint Surg Am 2001;83:1149-56.

[6] Mancuso CA, Charlson ME. Does Recollection Error Threaten the Validity of Cross-Sectional Studies of Effectiveness? Med Care 1995;33:AS77-88.

[7] Wilson J, Baker P, Rangan A. Is Retrospective Application of the Oxford Shoulder Score Valid? J Shoulder Elbow Surg 2009;18:577-80.

[8] Dawson EG, Kanim LE, Sra P, Dorey FJ, Goldstein TB, Delamarter RB, *et al*. Low Back Pain Recollection Versus Concurrent Accounts: Outcomes Analysis. Spine (Phila Pa 1976) 2002;27:984-93; discussion 994.

[9] Pellise F, Vidal X, Hernandez A, Cedraschi C, Bago J, Villanueva C. Reliability of Retrospective Clinical Data to Evaluate the Effectiveness of Lumbar Fusion in Chronic Low Back Pain. Spine (Phila Pa 1976) 2005;30:365-8.

[10] Liang KY, Zeger SL. Longitudinal Data Analysis Using Generalized Linear Models. Biometrika 1986;73:13-22.

[11] Zeger SL, Liang KY. Longitudinal Data Analysis for Discrete and Continuous Outcomes. Biometrics 1986;42:121-30.

[12] Vens M, Ziegler A. Generalized Estimating Equations and Regression Diagnostics for Longitudinal Controlled Clinical Trials: A Case Study. Comput Stat Data Anal 2012;56:1232-42.

[13] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria: 2018. Available from: https://www.R-project.org

[14] Vanegas LH, Rondón LM, Paula GA. Glmtoolbox: Set of Tools to Data Analysis using Generalized Linear. 2022. Available from: https://wwwCRAN.R-project.org/package=glmtoolbox

[15] Marsh J, Bryant D, MacDonald SJ. Older Patients can Accurately Recall their Preoperative Health Status Six Weeks Following Total Hip Arthroplasty. J Bone Joint Surg Am 2009;91:2827-37.

[16] Lindsay GM, Niven KA, Brodie EE, Gaw A, Belcher PR. Accuracy of Patient Recall of Preoperative Symptom Severity (Angina and Breathlessness) at One Year Following Aorta-Coronary Artery Bypass Grafting. J Clin Nurs 2009;18:418-25.

[17] Litwin MS, McGuigan KA. Accuracy of Recall in Health-Related Quality-of-Life Assessment Among Men Treated for Prostate Cancer. J Clin Oncol 1999;17:2882-8.

[18] Lowe JT, Li X, Fasulo SM, Testa EJ, Jawa A. Patients Recall Worse Preoperative Pain after Shoulder Arthroplasty than Originally Reported: A Study of Recall Accuracy using the American Shoulder and Elbow Surgeons score. J Shoulder Elbow Surg 2017;26:506-11.

[19] Pincus T, Pearce S, McClelland A, Turner-Stokes L. Self-Referential Selective Memory in Pain Patients. Br J Clin Psychol 1993;32:365-74.

[20] Randsborg PH, Adamec D, Cepeda NA, Ling DI. Two-Year Recall Bias after ACL Reconstruction is Affected by Clinical Result. JB JS Open Access 2021;6:e20.00164.

[21] Shetty T, Nguyen JT, Cogsil T, Tsiouris AJ, Niogi SN, Kim EU, *et al*. Clinical Findings in a Multicenter MRI Study of Mild TBI. Front Neurol 2018;9:836.

[22] Holden MP, Duff-Canning SJ, Hampson E. Sex Differences in the Weighting of Metric and Categorical Information in Spatial Location Memory. Psychol Res 2015;79:1-18.

[23] Schwartz CE, Sajobi TT, Lix LM, Quaranto BR, Finkelstein JA. Changing Values, Changing Outcomes: The Influence of Reprioritization Response Shift on Outcome Assessment after Spine Surgery. Qual Life Res

2013;22:2255-64.

[24] Herrmann D. Reporting Current, Past, and Changed Health Status. What we know about Distortion. Med Care 1995;33:AS89-94.

[25] McPhail S, Beller E, Haines T. Reference Bias: Presentation of Extreme Health States Prior to EQ-VAS Improves Health-Related Quality of Life Scores. A Randomised Cross-Over Trial. Health Qual Life Outcomes 2010;8:146.

[26] McPhail S, Haines T. Response Shift, Recall Bias and their Effect on Measuring Change in Health-Related Quality of Life amongst Older Hospital Patients. Health Qual Life Outcomes 2010;8:65.

[27] Tibbles AC, Waalen JK, Hains FC. Response Set Bias, Internal Consistency and Construct Validity of the Oswestry Low Back Pain Disability Questionnaire. J Can Chiropr Assoc 1998;42:141-9.

[28] Babul N, Darke AC, Johnson DH, Charron-Vincent K. Using Memory for Pain in Analgesic Research. Ann Pharmacother 1993;27:9-12.

[29] Singer AJ, Kowalska A, Thode HC Jr. Ability of Patients to Accurately Recall the Severity of Acute Painful Events. Acad Emerg Med 2001;8:292-5.

DOI: http://dx.doi.org/10.18053/jctres.08.202206.012

Journal of Clinical and Translational Research 2022; 8(6): 594-601

Journal of Clinical and Translational Research

Journal homepage: http://www.jctres.com/en/home

ORIGINAL ARTICLE

Accuracy of post-operative recall by degenerative cervical myelopathy patients using the modified Japanese Orthopaedic Association scale

Supplementary File

Modified Japanese Orthopaedic Association 17 Scoring Table (1994 version) (please mark "√" on the corresponding evaluation score)

Note:

When the selection is between two scores, the lower score shall be selected;

When the left and right nerve disorders are different, the lower score side score shall be recorded.

Scoring items	Scoring items
A. Sports function	**B. Sensory function**
I. Fingers: • Can't use any tableware, including chopsticks, spoons or forks, and/or fasten buttons (0 point) • Can't use chopsticks or write, can barely use spoons and knives and forks (1 point) • Can use chopsticks to hold large pieces of food, can barely write, can fasten large clothes (2 points) • Can use chopsticks, can't write words quickly, can fasten buttons (3 points) • Normal (4 points)	I. Pain: • Sensation of upper limbs completely disappears (0 point) • Only 50% or less of normal sensation and (or) severe pain and numbness (0.5 point) • Only 60% or less of normal sensation and (or) moderate pain and numbness (1 point) • Only slight numbness (normal touch) (1.5 point) • Normal (2 points)
III. Shoulder and elbow joints: Six grade muscle strength evaluation (MMT) I Muscle strength of trunk, deltoid muscle, and biceps brachii muscle, select the weaker one to record. • Deltoid muscle or biceps brachii muscle strength ≤ Grade 2 (2 points) • Deltoid muscle or biceps brachii muscle strength=Grade 3 (−1 points) • Deltoid muscle or biceps brachii muscle strength=Grade 4 (−0.5 points) • Deltoid muscle or biceps brachii muscle strength=Grade 5 (0 points)	II. Trunk: • The sense of tenderness disappears completely (0 point) • Only 50% or less of normal feeling and/or severe pain and numbness are completely disappeared (0.5 point) • Only 60% or less of normal feeling and/or moderate pain and numbness (1 point) • Only slight numbness (normal touch) (1.5 point) • Normal (2 points)
III. The lower limbs: • Cannot stand and walk independently (0 point) • Able to stand, but unable to walk (0.5 point) • Walking on the flat ground requires crutches or other supports (1 point) • Walking on flat ground without support, but with unstable gait (1.5 points) • Do not support when walking on the flat ground, but must grasp the handrail when climbing the stairs (2 points) • wAble to go upstairs by oneself, only when going downstairs must grasp the railings by hand (2.5 points) • Able to walk at a fast pace, but not fast, and has a clumsy gait (3 points) • Normal (4 points)	III. The lower limbs: • The sense of tenderness of lower limbs disappears completely (0 point) • Only 50% or less of normal feeling and (or) severe pain and numbness (0.5 point) • Only 60% or less of normal feeling and (or) moderate pain and numbness (1 point) • Only slight numbness (normal touch) (1.5 point) • Normal (2 points)

DOI: http://dx.doi.org/10.18053/jctres.08.202206.012