What is the correlation between patient-reported outcome measure (PROM) scores and patient satisfaction following elective reverse total shoulder replacement?

Rachael L. C. Daw¹, Jo Gibson¹, Denise Prescott²,
Laura Bonnett³ and Matthew Smith⁴

Abstract

Background: Outcomes of reverse total shoulder arthroplasty (RTSA) have typically been assessed using the same instruments as anatomical shoulder arthroplasties. However, to date, there has been a lack of investigation with respect to the correlation of such scores and patient satisfaction in the RTSA population.

Methods: The Oxford Shoulder Score (OSS) and Quick Disabilities of the Arm, Shoulder and Hand (QD) score were prospectively collected in 38 RTSA patients (41 shoulders) postoperatively. Scores were then evaluated to establish whether or not they correlated with patient satisfaction at a minimum of 1 year postoperatively.

Results: The correlation coefficient for the OSS and patient satisfaction was found to be 0.313 ($p = 0.011$) and the correlation coefficient for the QD score and patient satisfaction was -0.292 ($p = 0.017$), showing a statistically significant but moderately weak relationship between the OSS and QD scores with patient satisfaction ($p < 0.05$).

Conclusions: The present study showed no strongly significant relationship between patient-reported outcome measure (PROM) scores and patient satisfaction following elective RTSA. These findings emphasise the need to question the appropriateness of standard PROM scores for the assessment of outcome and success following elective RTSA.

Keywords

Patient-reported outcome measures, patient satisfaction, PROM scores, reverse total shoulder arthroplasty

Date received: 2nd May 2017; revised: 25th October 2017; accepted: 31st October 2017

Introduction

The treatment of degenerative and traumatic shoulder arthritis in the presence of a functional rotator cuff (RC) has been successfully addressed with the use of anatomical joint replacement.¹ In the absence of a functional RC, as in rotator cuff tear arthropathy (RCTA), there is insufficient soft tissue constraint to counteract the subluxing effect of the deltoid muscle to allow satisfactory shoulder function.^{2,3} Reversed total shoulder arthroplasty (RTSA) aims to re-centre the arc of rotation of the glenohumeral joint. Modern lateralized designs restore the glenoid–humeral relationship and the deltoid pulley improving stability and functional outcome but avoiding scapula notching that was problematic in the early prostheses.^{1–4} Clinical studies

evaluating the outcome of RTSA indicate that it has lived up to the biomechanical concept, with consistent successful restoration of range of movement (ROM) above 90° of elevation and significant pain relief.

¹Physiotherapy Department, Royal Liverpool Hospital, Prescot Street, Liverpool, UK

²Faculty of Health and Life Sciences, University of Liverpool, Liverpool, UK

³Department of Biostatistics, University of Liverpool, Liverpool, UK

⁴Department of Orthopaedics and Trauma, Royal Liverpool Hospital, Prescot Street, Liverpool, UK

Corresponding author:

Rachael L. C. Daw, Physiotherapy Department, Royal Liverpool Hospital, Prescot Street, Liverpool L7 8XP, UK.
Email: rachael.daw@rlbuht.nhs.uk

External and internal rotations do, however, remain limited, resulting in continued difficulty with some functional activities.^{2,4}

The ageing population⁶, coupled with increasing indications for RTSA, have meant that the procedure now accounts for a large proportion of all shoulder replacements performed in Europe and the USA.^{5,6,7} Irrespective of the indication for surgery the primary goal of RTSA is usually pain relief and, in many cases, ongoing functional limitation is both expected and accepted.⁸ Despite the significant differences in surgical procedure, prosthesis implant and expectation of outcome, RTSA is currently assessed in the same way as patients undergoing other types of shoulder arthroplasty surgery. Accordingly, questions have been raised regarding the suitability of this approach for this patient group.

Most national joint registries report the outcomes of arthroplasty surgery in terms of revision rates or formal survival analyses and performance is based on revision as the endpoint.^{8,9} Revision rate alone has, however, been demonstrated not to accurately reflect the success of surgery.^{8,9} Methods to define a successful outcome following shoulder arthroplasty are subject to a degree of contention¹⁰ and increasing emphasis on the patient's perspective has meant that they have had to evolve to encompass a broader definition of a satisfactory result.^{8,11}

Patient-reported outcome measures (PROMs) are now commonly used in many healthcare settings,¹² and are often a stipulated requirement for healthcare commissioners. In the case of shoulder surgery, region-specific PROMs such as the Oxford Shoulder Score^{13,14} and Quick Disabilities of the Arm, Shoulder and Hand (QD)¹⁵ score are commonly used alongside generic measures of satisfaction and health status. The degree to which subjective PROMs reflect objective findings and how well subjective and objective measures reflect patient satisfaction is yet to be established.^{8,10}

PROMs questionnaires do not ask about a patient's satisfaction with healthcare or seek their opinions about how successful their treatment was.¹⁶ Some patients may continue to have functional limitations following surgery, which is reflected in a 'poor' PROM score. However, a patient's expectation and acceptance of chronic functional limitations might mean that they are satisfied with the outcome and their surgery has been more successful than the PROM scores would suggest.⁸ Patient satisfaction can provide the ultimate endpoint to healthcare quality and is thus an essential part of quality assessment.¹⁶⁻¹⁸

Study aim

The aim of the present study was to evaluate the relationship between PROM scores and patient-reported

satisfaction following elective reverse total shoulder arthroplasty.

Null hypothesis: there will be no positive relationship between PROM scores and patient-reported satisfaction scores in patients following elective RTSA.

Materials and methods

To address the study aims, a non-experimental fixed design was used to measure the relationship between the variables OSS, QD scores and patient satisfaction scores (measured on a numerical rating scale; NRS). To explore the relationship between the variables, data routinely collected at patients' annual follow-up appointments were reviewed and then subjected to correlational statistics. The study evaluated outcomes of an existing service without a reference standard and analyzed data regarding an established intervention.¹⁹

The study used a sample of convenience, taking data from a consecutive patient cohort who all underwent RTSA surgery. When completing PROM scores, patients were asked to complete a consent statement.

Using National Research Ethics Service (NRES) guidance,¹⁹ it was established that the study would be a service review and NRES approval was therefore not necessary. The study was registered with and approved by the local audit office, who granted permission to go ahead.

Data were routinely collected from consecutive patients who had shoulder arthroplasty surgery. Patients of four senior orthopaedic surgeons using the same Exactech Equinox prosthesis system (Exactech UK Ltd, Redditch, UK) were included. These patients were routinely reviewed annually in a dedicated shoulder arthroplasty review clinic, around the anniversary of their procedure, and standardized outcome scores were completed (OSS, QD score and patient satisfaction NRS).

Patients were included if they had had a primary elective RTSR with greater than 12 months of follow-up, and were over 65 years of age at the time of surgery. Patients who had revision procedures, were under 65 years, had not consented for their data to be used, or were not capable of proving consent, were excluded from the study, along with those who had procedures for acute trauma.

To establish the final data set for the study, all of the existing data from the RTSA patient cohort, which had been collected from the commencement of the shoulder arthroplasty review clinic in February 2012 until February 2016, were reviewed. Individual patient datasets that satisfied all of the inclusion criteria and also included a complete set of the required outcome measures at a single point in time were extracted from the database.

Statistical analysis

The data were subjected to statistical analysis to establish whether a significant correlational relationship existed between either of the absolute postoperative PROM scores (OSS or QD score) and patient satisfaction score.

Statistical analysis of the data, aiming to establish the strength and direction of any correlation between the variables, was completed with the guidance of a statistician from the University of Liverpool. The normality of numeric distribution of the three scores (satisfaction, OSS and QD score) was examined via histograms. The scores were all found not to be normally distributed and were therefore subjected to various transformations, such as log and square root (QD score) and quadratic and cube (OSS and patient satisfaction), aiming to determine whether normality could be established.^{14,19} Normality could not be established in any of the scores despite the transformations; therefore, it was not possible to use the standard Pearson's correlation coefficient for parametric data. Kendall's tau-b correlation coefficient, which does not assume normal distribution of the data, was calculated to examine the relationships between OS and patient satisfaction and between QD score and patient satisfaction. Kendall's tau-b was chosen in preference to the more standard Spearman's rank correlation coefficient as a result of the large number of tied ranks.²⁰

In the case of all coefficients, a *p*-value was also calculated to determine statistical significance, $p < 0.05$ was considered as statistically significant. This level was selected because, in the small sample, this is likely to represent a significant change.^{21,22}

Results

Sixty-seven patients were identified who were more than 12 months post elective RTSA. Twelve patients were not eligible for the study because they had undergone their surgery to revise a previous arthroplasty procedure and a further seven were excluded because they were under 65 years of age at the time of surgery. Of the remaining 48 patients, one had refused to attend for postoperative follow-up, three had deceased prior to the 12-month assessment and one had not yet attended for follow-up. Five patients did not have a complete set of outcome scores. Of the remaining 38 patients, three underwent bilateral RTSA. Because each shoulder was assessed and scored separately, both sets of data for these patients were included in the study. Complete data sets were available for 85% (41/48) of all eligible procedures.

The indications for surgery according to diagnosis of the included patients are shown in Table 1.

Table 1. Indications for surgery.

Indication for surgery	Number of patients (%)
RCTA	26 (63)
Rotator cuff tear without OA	5 (12)
OA (without cuff tear)	6 (15)
Post-traumatic OA	3 (7.5)
Chronic dislocation	1 (2.5)

RCTA, rotator cuff tear arthropathy; OA, osteoarthritis.

The age range for patients at the time of surgery was from 65 years to 87 years (mean 74 years, median 73 years). Sixteen procedures (39%) were undertaken on males and 25 (61%) were undertaken on females. The timescale for the postoperative review assessment ranged from 12 months to 37 months postoperatively (mean 24 months).

The OSS ranged from 16 to 48 [mean (SD) 36.63 (9.8)]. The QD score ranged from 0 to 70.55 [mean (SD) 29.66 (20.84)]. The patient satisfaction scores ranged from 6 to 10 [mean (SD) 9.02 (1.172)]. Patient satisfaction was scored on an NRS, with 0 representing completely unsatisfied and 10 representing completely satisfied.

Using the Kendall-tau B, the correlation coefficient for the OSS and patient satisfaction was found to be 0.313 ($p = 0.011$) and the correlation coefficient for the QD score and patient satisfaction was -0.292 ($p = 0.017$). There is a statistically significant relationship between the OSS ($p < 0.05$) and patient satisfaction, which is moderately weak in the positive direction. The relationship between the QD score and patient satisfaction is also statistically significant ($p < 0.05$) but moderately weak in the negative direction.

Limitations and strengths

The present study included a sample of convenience; nonetheless, the employment of robust inclusion and exclusion criteria assured a clearly defined cohort. As such, the study findings are applicable to this group and relevant to the local population; however, because of the sample selection method, they cannot be assumed to more widely represent outcomes for other patients following RTSA.²³

The inclusion of patients with variable indications for surgery and comorbidities could confound the clarity of the findings. The small sample size may affect the accuracy of the study findings and preclude comparison with the broader population. The relatively short-term

mean follow-up timescale also limits interpretation of the results in relation to long-term outcomes. Limiting the study to one treatment centre and one type of prosthesis has ensured a degree of consistency within the group; on the other hand, this could be considered as a further confounding factor when extrapolating findings to the wider patient population and we do not know how this particular prosthesis compares to others.

The NRS, as used to measure patient satisfaction in this study, has been validated for reporting the outcome of patients with various shoulder disorders as part of the Penn Shoulder Score; it was also found to have sufficient reliability and validity to be used in isolation.²⁴ Commonly used in clinical practice and research settings, the NRS is easily understood and quick to complete, resulting in high response rates. However, the NRS has been criticised for being too simplistic and not able to accurately reflect the multifaceted nature of satisfaction.²⁵ It must be considered, therefore, that the limited relationship demonstrated between the PROM and satisfaction scores in the present study could partly be the result of an erroneous estimation of satisfaction based on an imperfect scoring system.

Standardized pre-operative data were not available for this entire patient cohort. Therefore, it was not possible to compare change in the scores with satisfaction. The addition of a general health score such as the EuroQol or SF-12 may also have added strength to the evaluation of patient outcome.^{26,27}

Discussion

The aim of the present study was to evaluate the relationship between commonly used PROM scores and patient-reported satisfaction following elective RTSA, with a view to increasing understanding of how these assessment methods reflect patient perceived outcome.

Currently, there are no published data available regarding the expected range for outcome scores such as the OSS and QD score following RTSA. The mean OSS for the study cohort was 36.63, falling notably short when compared to the mean OSS (45.2) of an asymptomatic population over 65 years of age.²⁸ Similarly the mean QD score of 29.66 appears to be considerably inferior to that demonstrated in the general population over 60 years of age (mean 18.9).²⁹ Interpreting the mean PROM scores found in the present study by comparison with the general population reference could infer that patient perceived outcome was poor. However, when considering the mean satisfaction score of the study population (9.02/10), an entirely contrasting and superior impression of success is deduced. This suggests that the PROM scores do not accurately reflect the success of the procedure as perceived by the patient.

The apparent disparity between PROM scores and satisfaction is reinforced by the results of correlation analysis. Although statistically significant, the correlations between both OSS and QD score and patient satisfaction were only moderately weak. These results suggest that the patient's perception of outcome represented by their own judgement of satisfaction does not closely align with that represented by the PROM scores. The implication of these findings is that these commonly used PROM scores may not be an appropriate method of outcome evaluation for this particular patient cohort. This highlights the need for further investigation into the most appropriate methods for assessing surgical success for RTSA.

The OSS and QD scores, both frequently used to assess RTSA outcomes, ask questions about general functional activities including tasks that are often impossible for patients with the pain and weakness associated with severe degenerative shoulder joint disease. Regardless of improvement in pain and ROM following RTSA, these tasks may continue to be impossible or only possible with some degree of difficulty because the amount of strength and movement required may not be achievable. Comorbidities that are not addressed by the surgical procedure would also potentially continue to exert the same limitations on functional ability postoperatively as they did pre-operatively. In the presence of ongoing functional difficulties, a patient would still not necessarily achieve a 'good' final outcome score, and possibly not have a great improvement in a score too; despite this, they may still have what they consider to be a very satisfactory outcome. The PROM scores in this situation would fail to provide accurate information regarding the patient's interpretation of the surgical outcome. These functional factors may account for the lack of correlation between PROMs and patient satisfaction. Future investigation of the correlation between change in PROMs (from pre- to postoperatively) and satisfaction may add valuable insight to these considerations.

Regarding the limitations in the application of PROM scores, it is also important to question the use of other outcome measures, such as patient satisfaction scores, which may well be disadvantaged by their own inadequacies.¹¹ Although it has been suggested that patient satisfaction is the ultimate objective of all healthcare intervention,^{13,16} there is little consensus on the best method for measuring it. To date, there is a lack of evidence for measuring satisfaction with orthopaedic intervention and no evidence to support any one method for measuring satisfaction in relation to shoulder arthroplasty surgery specifically. Nonetheless it could be argued that the NRS currently represents the most effective method for assessing satisfaction

(but not overall outcome) with shoulder surgery and was the best available option for the present study. The NRS has been employed in recent literature examining the relationships between the various outcomes of shoulder arthroplasty,^{11,30} reinforcing its relevance for utilisation in the present study. More recently, Van Berckel et al.³¹ have also found a similarly weak correlation between the QD score and an NRS score of patient satisfaction. (Their study differed from the present study, however, in that the NRS was used to evaluate patient satisfaction with overall management rather than outcome)³¹.

The findings of the present study, which suggest a limited relationship between PROM scores and patient satisfaction, open a debate regarding the relevance of existing scores to RTSA. The present study is the first to provide specific data examining the correlation between PROMs and patient satisfaction in the RTSA population. As a result of its limitations, the present study is unable to answer all of the pertinent questions, although it highlights issues for consideration. Further studies of this type are required to firmly establish whether existing outcome measures adequately reflect the patient's perception and satisfaction with surgery. Prospective studies including larger patient cohorts, longer term follow-up, different age groups and indications for surgery would facilitate further clarification of the argument regarding the appropriateness of PROM scores for the assessment of RTSA.

In the current clinical and research contexts, the present study highlights the need to take great care when interpreting clinical scores and when planning studies of treatment effects. The routine use of PROMS in healthcare settings must be considered carefully if meaningful information is to be disseminated to patients and clinicians. There is an important risk with some patient groups, such as those undergoing RTSA, with respect to outcomes data being liable to misinterpretation, resulting in misinformation, if the full context of expectations and outcomes is not properly understood.

Conclusions

The results of the present study demonstrate that there is no strong and significant relationship between PROM scores and patient satisfaction in a cohort of patients following RTSA. These findings suggest that, in the future, PROM scores used for evaluating RTSA should be interpreted with caution. Further investigation is required to establish the methods of assessment that most accurately reflect the patient experience and offer a true representation of value and outcome.

Acknowledgements

The authors would like to acknowledge and thank the Consultant Orthopaedic Surgeons who operated on patients included in the present study: Mr Matthew Smith, Mr Matthew Kent, Mr Inigo Guisasaola and Mr Peter Brownson.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: The physiotherapy research fellowship held by Mrs Daw is financially supported by Exactech UK Ltd. Mr Matthew Smith holds a consulting contract with Exactech UK.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Exactech UK Ltd and Arthritis UK.

Ethical Review and Patient Consent

Not required for this article.

References

1. Jazayeri R and Kwon YW. Evolution of the reverse total shoulder prosthesis. *Bull NYU Hosp Jt Dis* 2011; 69: 50–55.
2. Baulot E, Sivereaux F and Boileau P. Grammont's idea: the story of Pauls Grammont's functional surgery concept and the development of the reverse principle. *Clin Orthop Relat Res* 2011; 469: 2425–2431.
3. Lee LH and Desai A. Reverse polarity shoulder replacement: current concepts and review of literature. *World J Orthop* 2014; 5: 255–261.
4. Boileau P, Watkinson DJ, Hatzidakis AM and Balg F. Grammont reverse prosthesis: design, rationale and biomechanics. *J Shoulder Elbow Surg* 2005; 14: 147S–161S.
5. Smith CD, Guyver P and Bunker TD. Indications for reverse shoulder replacement. A systematic review. *J Bone Joint Surg* 2012; 94B: 577–583.
6. United Nations, Department of Economic and Social Affairs, Population Division (2013). World Population Prospects: The 2012 Revision. http://esa.un.org/unpd/wpp/publications/Files/WPP2012_HIGHLIGHTS.pdf Accessed 01/12/2015.
7. Schairer WW, Nwachukwu BU, Lyman S, Craig EV, Gulotta LV. National utilization of reverse total shoulder arthroplasty in the United States. *J Shoulder Elbow Surg*. 2015; 24: 91–97 National Joint Registry. 12th Annual Report. 2015. <http://www.njrcentre.org.uk/njrcentre/Portals/0/Documents/England/Reports/12th%20annual%20report/NJR%20Online%20Annual%20Report%202015.pdf>.
8. Rees JL, Dawson J, Hand GCR, et al. The use of patient-reported outcome measures and patient satisfaction ratings to assess outcome in hemi-arthroplasty of the shoulder. *J Bone Joint Surg* 2010; 92B: 1107–1111.

9. Murray DW and Frost SJ. Pain in the assessment of total knee replacement. *J Bone Joint Surg* 1998; 80: 426–431.
10. Harreld K, Clark R, Downes K, Virani N and Frankle M. Correlation of subjective and objective measures before and after shoulder arthroplasty. *Orthopaedics* 2013; 36: 808–814.
11. Zywiell MG, Mahomed A, Gandhi R, Peruccio AV and Mahomed NN. Measuring expectations in orthopaedic surgery: A systematic review. *Clin Orthop Relat Res* 2013; 471: 3446–3456.
12. Gibbons E, Fitzpatrick R. Patient reported outcome measures: Their role in measuring and improving patient experience. 2012. patientexperience.org/article/patient-reported-outcome-measures-their-role-in-measuring-and-improving-patient-experience (Accessed 23/11/2015).
13. Dawson J, Fitzpatrick R and Carr A. Questionnaire in the perceptions of patients about shoulder surgery. *J Bone Joint Surg* 1996; 78B: 593–600.
14. Dawson J, Rogers K, Fitzpatrick R and Carr A. The Oxford Shoulder Score revisited. *Arch Orthop Trauma Surg* 2009; 129: 119–123.
15. Beaton DE, Wright JG, Katz JN. Development of the QuickDASH: Comparison of three item approaches. *J Bone Joint Surg (Am)*. 2005; 87-A: 1038–1046.
16. Devlin NJ, Appleby J. Getting the most out of PROMS, putting health outcomes at the heart of NHS decision making. The Kings Fund. London; 2010.
17. Chow A, Mayer EK, Darzi AW and Athanasiou T. Patient reported outcome measures: The importance of patient satisfaction in surgery. *Surgery* 2009; 146: 435–443.
18. Brokelman RBG, Haverkamp D, van Loon C, Hol A, van Kampen A and Veth R. The validation of the visual analogue scale for patient satisfaction after total hip arthroplasty. *Eur Orthop Traumatol* 2012; 3: 101–105.
19. National Patient Safety Agency. Defining research: NRES guidance to help you decide if your project requires review by a Research Ethics Committee. 2009.
20. Altman G, Moore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. In: Altman G, Machin D, Bryant TN, Gardner MJ (Eds) *Statistics with Confidence* (2nd edition). Bristol. BMJ Books. 2000. p186–187.
21. Altman G and Gardner MJ. Means and their differences. In Altman G, Machin D, Bryant TN, Gardner MJ (Eds) *Statistics with Confidence* (2nd edition). Bristol. BMJ Books. 2000. p32–33.
22. Banerjee A. Medical Statistics made clear: An introduction to basic concepts. London. Royal Society of Medicine Press. 2003.
23. Robson C. Real World Research: a resource for users of social research methods in applied settings (3rd edition). Chichester: John Wiley & Sons Ltd, 2011.
24. Leggin BG, Michener LA, Shaffer MA, Brenneman SK, Iannotti JP and Williams GR. The Penn Shoulder Score: Reliability and Validity. *J Orthop Sports Phys Ther* 2006; 36: 138–151.
25. Lloyd H, Jankinson C, Hadi M, Gibbons E and Fitzpatrick R. Patient reports of the outcome of treatment: a structured review of approaches. *Health Qual Life Outcomes* 2014; 12: 1–9.
26. Chen AL, Bain EB, Horan MP, Hawkins RJ, Littleton NH and Vail CO. Determinants of patient satisfaction with outcome after shoulder arthroplasty. *J Shoulder Elbow Surg* 2007; 16: 25–30.
27. Younis F, Sultan J and Hughes PJ. The range of the Oxford Shoulder Score in the asymptomatic population: a marker for post-operative improvement. *Ann R Coll Surg Eng* 2011; 93: 629–633.
28. Aasheim T and Finsen V. The DASH and the QuickDASH instruments in the general population in Norway. *J Hand Surg* 2014; 39E: 140–144.
29. Van Berckel MMG, Bosma NH, Hageman MGJS, Ring D and Vranceanu A-M. The correlation between a numerical rating scale of patient satisfaction with current management of an upper extremity disorder and a general measure of satisfaction with the medical visit. *Hand* 2017; 12: 202–206.