

A study to assess patient-reported outcome measures (PROMs) and to investigate the practicality of using PROMs in a surgical office

“If you cannot measure it, you cannot improve it.”

—Sir William Thomson, Lord Kelvin

ABSTRACT

Background: After a Ministry of Health conference on outcome analysis, a study was proposed to explore the value of incorporating patient-reported outcome measures into the daily activity of a surgical office.

Methods: Subjective survey and objective clinical testing methods were used. Patients undergoing hip or knee total joint replacement at a nonacademic centre were asked to complete outcome questionnaires before and after surgery and the results were compared with clinical assessments by a physiotherapist. Most hip and knee patients had primary joint replacement, but some had revision surgery. A small cohort of patients undergoing shoulder surgery for rotator cuff repair was studied as well. The relationships between the subjective and objective test results were analyzed using a 2-tailed Pearson correlation test. The relationships identified were then displayed in scatter plot graphs, with each circle on the graph

indicating the data comparing one test with another for one patient. The data were also subjected to further statistical analysis.

Results: The subjective patient-reported results for total joint replacements correlated closely with the objective clinical assessments, indicating that patient-reported results are sufficient on their own to show outcomes. There was a correlation between results from two subjective tests for rotator cuff repairs. However, no correlation was found when results from subjective and objective tests for shoulder patients were compared. The evaluation concept of “minimum clinically important difference” was found to be useful and more meaningful when assessing results for the hip and knee patients.

Conclusions: The use of patient-reported outcome measures with values for minimum clinically important difference is recommended for evaluating outcomes and assessing appropriateness for any surgical or medical intervention. Patient-

reported outcome measures can also assist clinicians in their self-assessment and be an adjunct to their continuing professional development. The province should take a leading role in implementing the use of patient-reported outcome measurement in the offices of surgeons and other practitioners. Implementation would involve helping with statistical analysis and establishing ways to obtain consent and handle collection and sharing of patient data to ensure privacy.

Dr Stanger is an orthopaedic surgeon practising with RebalanceMD, a team of physicians and adjunct health professionals serving patients in the Vancouver Island Health Authority. Ms Christine Morrison is a physiotherapist and former project manager for patient-reported outcome measures at Island Health. Dr Yazganoglu is a physician and former senior analyst for the Surgical Patient Registry of the Provincial Health Services Authority. Dr Bhalla is a public health administrator and former director of Surgical and Anaesthetic Services, Health Services Policy and Quality Assurance Division, Ministry of Health.

This article has been peer reviewed.

Background

McGrail and colleagues¹ have written recently that routine measurement of patient-reported outcomes should be instituted in Canadian health care. Although numerous studies have considered the use of patient-reported outcome measures (PROMs) and related them to clinical tests,²⁻⁵ few of these studies have originated in a non-academic setting.⁶

In December 2010, the BC Ministry of Health sponsored a symposium on patient-reported outcome studies. As a result of that meeting, Margi Bhalla and Michael Stanger approached the Specialist Services Committee (a Doctors of BC and Ministry of Health joint committee) to fund a study of outcome measurement for patients attending a joint replacement clinic (JRC) operated by Vancouver Island Health Authority (Island Health) as a way to explore incorporating PROMs into the daily activity of a surgical office.

The objectives of the study proposed by Bhalla and Stanger were to record the outcomes of joint replacement surgery and to assess the practicality of undertaking outcome studies in a nonacademic, routine surgical setting. To that end, the study also included a small cohort of patients having shoulder surgery for rotator cuff repair. These patients were managed solely through a surgical office rather than through the JRC.

Methods

Ultimately, the study included 1231 hip and knee patients referred to the joint replacement clinic in Victoria, BC, and 39 shoulder patients undergoing rotator cuff repair. Most hip and knee patients had primary joint replacement, but some had revision surgery.

The patients who had been booked for hip or knee replacement were asked to sign a release to use their in-

formation, and approval was sought from the Island Health ethics board, since the health authority was responsible for the clinic where testing was initiated.

In the case of hip and knee patients, testing was done before each patient's surgery and at 3 to 6 months after surgery. Final assessment was done at 12 months after surgery based on support for the 1-year interval in an article by Browne and colleagues.⁷ In the case of shoulder patients, the final assessment was at 6 months, since that seemed to coincide with the patient's recovery.

The principal investigators were Margi Bhalla, representing the Ministry of Health, and Michael Stanger, a retired orthopaedic surgeon associated with the Section of Orthopaedics and RebalanceMD in Victoria (the surgical office for the study). The project manager was Christine Morrison, a physiotherapist responsible for collecting all the data as well as for performing the clinical tests that were used to determine the accuracy of PROMs evaluations. Data collection and collation were managed by Erdem Yazganoglu, a physician with the Surgical Patient Registry, who ensured the meaningful presentation of information from the large amount of data collected.

Subjective tests

Two types of subjective patient-reported outcome tests were used in the study. One was a patient-completed questionnaire about general health status and the other was a test specific to the patient's condition.

The general health status instrument used to evaluate all patients was the UK version of the EuroQol 5D-SL (EQ-5D), which consists of five general questions as well as the EQ-5D visual analog scale (EQ-VAS) that patients are asked to rate themselves on.

Specific conditions were assessed using the Oxford hip score (OHS), Oxford knee score (OKS), and the Oxford shoulder score (OSS). Each of these tests asks the patient 12 questions related to the joint of concern and allows the patient to select answers describing varying degrees of functional impairment.

Routine measurement of patient-reported outcomes should be instituted in Canadian health care.

The maximum score for the EQ-5D is 1.0, while the maximum for the EQ-VAS is 100 and the maximum for each of the Oxford tests is 48.

Patients filled out paper forms for the EQ-5D, EQ-VAS, and the Oxford tests. These are available for completion electronically, but study funds did not permit having patients input the information electronically.

Objective tests

Standardized tests of mobility and activity specific to the joint in question were conducted by a qualified physiotherapist. Tests used were the Harris hip score (HHS), the Knee Society score (KSS), and the Constant shoulder score (CSS).

The Harris hip test and the Knee Society test are both scored out of 100, with a higher score indicating better function. The Constant shoulder test is also scored out of 100, but improvement is indicated by a reduction rather than an increase in score.

Preoperative clinical testing was done at the JRC and postoperative clinical testing was done at the Re-balanceMD clinic, often at the time of a follow-up appointment with the surgeon.

Correlations sought

The relationships between the subjective and objective test results were analyzed using a 2-tailed Pearson correlation test. The relationships identified were then displayed in scatter

plot graphs, with each circle on the graph indicating the data comparing one test with another for one patient. The data were also subjected to further statistical analysis.

Results

The graphs displaying results for the hip patients (Figure 1 and Figure 2) and for the knee patients (Figure 3) each show a well-aligned distribution of data points indicating a strong correlation between the results for the

two tests being compared, and statistical calculations show that the relationships are statistically significant.

Despite the fact that fewer shoulder patients were assessed, the graph for results from two subjective tests at 3 months (Figure 4) shows a statistically significant relationship. However, when results from the subjective and objective tests for shoulder patients are compared (Figure 5), no correlation is seen.

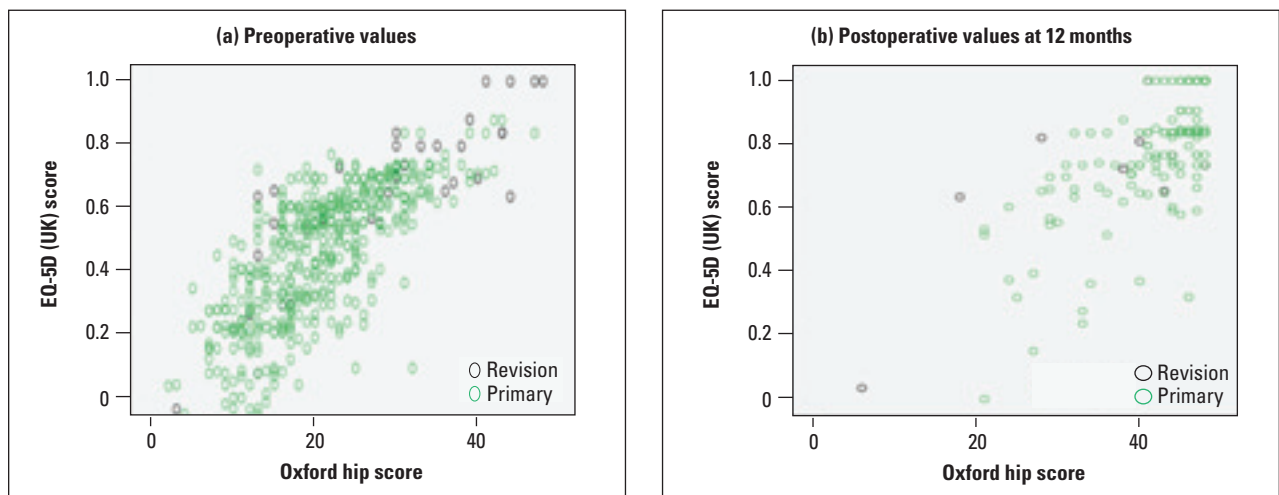


Figure 1. Correlations between two subjective hip tests: Comparison of EO-5D and Oxford hip score values before (a) and after (b) surgery.

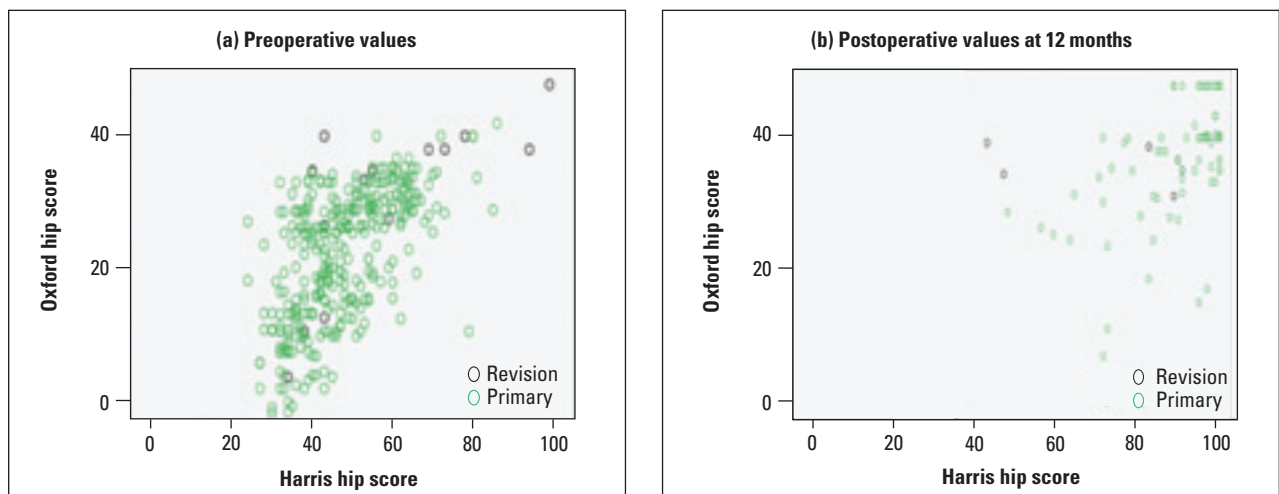


Figure 2. Correlations between a subjective hip test and an objective hip test: Comparison of Oxford hip score and Harris hip score values before (a) and after (b) surgery.

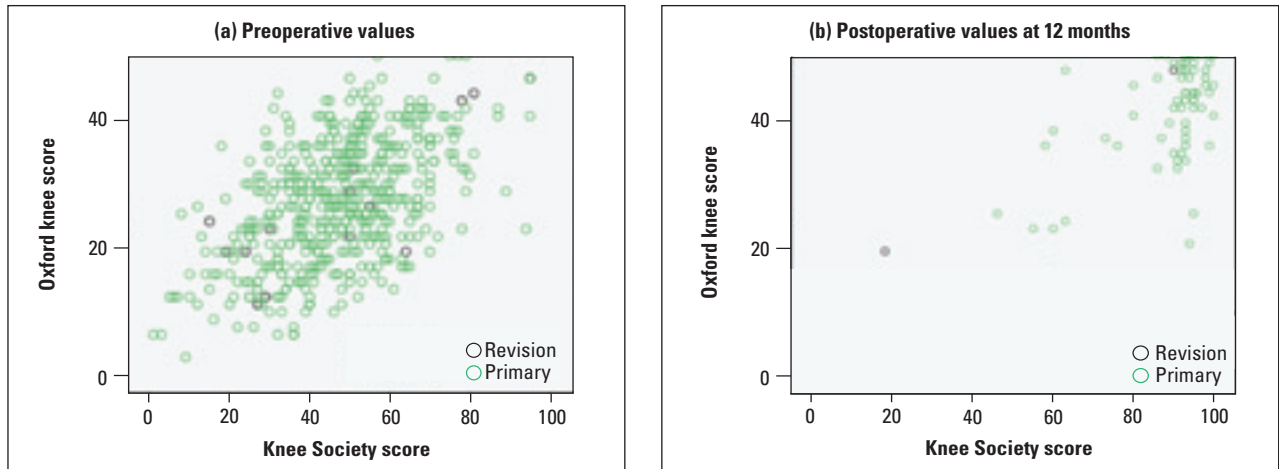


Figure 3. Correlations between a subjective knee test and an objective knee test: Comparison of Oxford knee score and Knee Society score values before (a) and after (b) surgery.

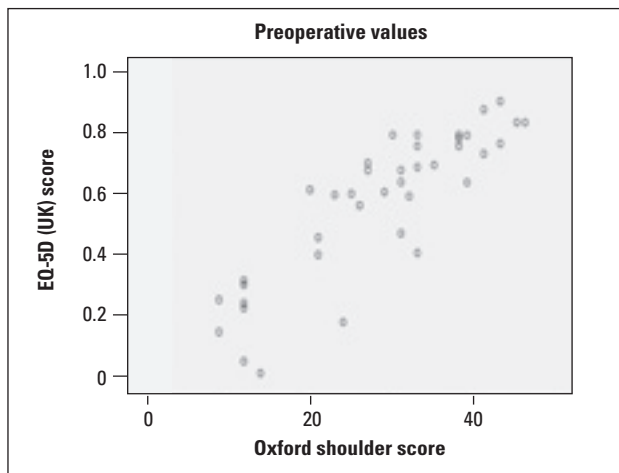


Figure 4. Correlations between two subjective shoulder tests: Comparison of EQ-5D score and Oxford shoulder score values before surgery.

The graphs displaying results for hip ... and knee patients ... indicate a strong correlation ... for the two tests being compared.

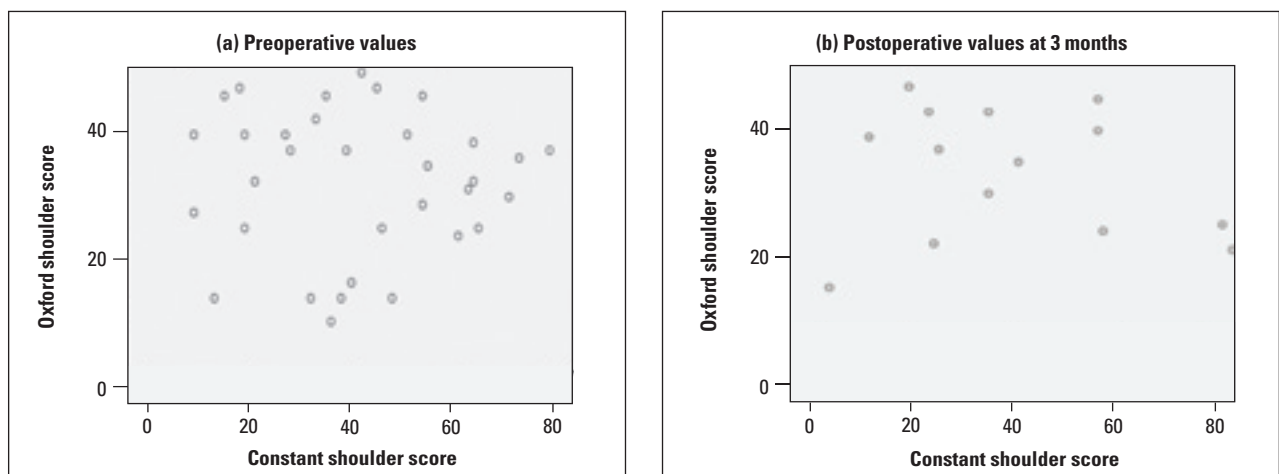


Figure 5. Correlations between a subjective and an objective shoulder test: Comparison of Oxford shoulder score and Constant shoulder score values before (a) and after (b) surgery.

Hip and knee patients

Based on patient responses to general and specific health indices, improvement can be seen for both total hip patients (**Table 1**) and total knee patients (**Table 2**) at 3 to 6 months post-operatively and at 12 months post-operatively. For hip patients, Judge and colleagues⁸ have determined a high level of patient satisfaction is associated with an Oxford hip score change of more than 14 points. In our study this change was 21.23 points. For knee patients, Judge and colleagues have determined the critical change in value for the Oxford knee score is 11 points. The change for knee patients in our study was 14.85 points.

Parker and colleagues⁹ have discussed the need to establish the smallest amount an outcome must change to be valued by patients and reflected in their responses on a PROMs ques-

tionnaire: “A shortcoming of these questionnaires is that their numerical scores lack a direct meaning or clinical significance. Because of this, the concept of the minimum clinically important difference (MCID) has been put forth as a measure for the critical threshold needed to achieve treatment effectiveness.”⁹ An acceptable number of patients reaching an MCID would “imply clinical significance and justification for implementation into clinical practice.”⁹

We defined the MCID value for our study as an improvement in Oxford scores of 14 points for hip patients and 11 points for knee patients. We found that 80.5% of hip patients and 68.9% of knee patients achieved this minimum clinically important difference by 12 months (set in bold in **Table 3**). We also found significant improvement when compar-

ing the 3-to-6 month results with the 12-month results for hips (61.7% versus 80.5%) and for knees (54.0% versus 68.9%). This supports using a 12-month follow-up interval for joint replacement assessments.

Shoulder patients

There were only 39 rotator cuff repair patients entered in the study (**Table 4**). These patients were much younger than the joint replacement patients and tended to be still working. During the study period many more patients had rotator cuff repairs but were not entered in the study, illustrating the difficulty of obtaining patient information if subjective testing is not an integral part of the patient documentation process and history taking.

The other issue found with this small cohort was the correlation between the different tests. The objective test that was selected did not correlate with the subjective tests. There were sufficient subjects to allow an evaluation at the outset, but too few subjects for evaluation at 6 months.

Including the shoulder group, however, was very instructive in terms of identifying the need to have the PROMs test as an integrated part of patient intake, pre-op assessment, and post-op follow-up.

Conclusions

Routine evaluation of outcomes is in the best interests of the patient. The patient-reported outcome test is a valid and reliable way to determine the results of specific treatments, and may be used on its own without clinical testing. Most PROMs tests have undergone validity testing, but in situations where a PROMs test has not been validated it may be necessary to use a clinical test as well until validity is established and the clinical test can be discontinued.

Table 1. Improvement measured for 546 total hip replacement patients.

Test	Pre-op values	Post-op values at 3-6 months	Change from pre-op to post-op values at 3-6 months	Post-op values at 12 months	Change from pre-op to post-op values at 12 months
EQ-5D (max 1.0)	0.43	0.74	+0.31	0.73	+0.30
EQ-VAS (max 100)	65.04	79.75	+14.71	78.88	+13.84
OHS (max 48)	20.32	37.20	+16.88	41.55	+21.23
HSS (max 100)	47.15	82.00	+34.85	89.66	+42.51

Table 2. Improvement measured for 680 total knee replacement patients.

Test	Pre-op values	Post-op values at 3-6 months	Change from pre-op to post-op values at 3-6 months	Post-op values at 12 months	Change from pre-op to post-op values at 12 months
EQ-5D (max 1.0)	0.53	0.73	+0.20	0.68	+0.15
EQ-VAS (max 100)	70.27	78.00	+7.73	77.53	+7.25
OHS (max 48)	22.87	34.10	+11.23	37.72	+14.85
KSS (max 100)	47.13	82.15	+35.02	90.58	+43.45

Table 3. Hip and knee replacement patients reaching minimum clinically important difference (MCID) at 3 to 6 months and at 12 months.

Measure	Hip replacement		Hip total	Knee replacement		Knee total
	Primary	Revision		Primary	Revision	
Difference at 3–6 months						
Joint replacements at or above MCID*	177	3	180	209	4	213
Assessments where difference can be calculated	287	16	303	383	8	391
Percentage of joint replacements at or above MCID*	61.7%	18.8%	59.4%	54.0%	50.0%	54.5%
Difference at 12 months						
Joint replacements at or above MCID*	120	3	123	144	6	150
Assessments where difference can be calculated	149	11	160	209	7	216
Percentage of joint replacements at or above MCID*	80.5%[†]	27.3%	76.9%	68.9%[†]	85.7%	69.4%

*Defined as a 14-point change in Oxford hip score and 11-point change in Oxford knee score

[†]Bold for emphasis

Relationship of PROMs to appropriateness

The use of patient-reported outcome measurement before and after surgery provides a clear indication of the outcome of the surgery. Patients themselves are registering whether they have improved as a result of the procedure. In addition, when patients register improvement this provides de facto evidence of the appropriateness of the intervention being studied.

Wright and colleagues¹⁰ have stated that variations in utilization rates for elective surgical procedures and the lack of systematic outcome evaluation can lead to concerns about the appropriateness of a procedure, and that this could be avoided by routinely evaluating indications and outcomes.

de Boer and colleagues¹¹ have investigated the relationship between the surgeon’s evaluation of the outcome and the patient’s evaluation. Their conclusion was that “The physician was more optimistic about the outcome of the operation than was justified according to the answers to a patient self-assessment questionnaire.”¹¹ In contrast, a report by Harris and colleagues¹² found a discrepancy of only 4% between the surgeon-reported outcome and the patient-

Table 4. Improvement measured for 39 rotator cuff repair patients.

Test	Pre-op values	Post-op values at 3 months	Change* from pre-op to post-op values at 3 months	Post-op values at 6 months	Change* from pre-op to post-op values at 6 months
EQ-5D (max 1.0)	0.57	0.59	+0.02	0.74	+0.16
EQ-VAS (max 100)	69.51	69.00	+0.51	84.55	+15.04
OSS (max 48)	28.23	30.90	+2.67	37.55	+9.32
CSS (max 100)	41.24	39.57	-1.67	29.80	-11.44

*Improvement is shown in two different ways: for the EQ-5D, EQ-VAS, and OSS tests, an increase in values indicates improvement; for the CSS test, a reduction in values indicates improvement.

reported assessment. Bream and colleagues¹³ also found a reasonable correlation between the physician’s evaluation and the patient’s.

The recent provincial interest in credentialing of practitioners may also relate directly to the use of PROMs on a routine basis as a way to ensure practitioners are following their own results in an organized and understandable fashion. The use and review of tests such as the EQ-5D and Oxford hip score relates directly to continued professional development.

Costs of PROMs

The costs related to using PROMs are difficult to assess because of a num-

ber of factors. In our study, much of the cost was for the clinical testing done to determine the accuracy of PROMs values, and more staff time was required because of a complicated method used to test individual patients and examine and record the related data. Costs would fall dramatically if patient-reported outcome measurement was incorporated automatically into surgical office procedures and if the independent clinical exam was not required.

Implications for the surgical office

This study shows that PROMs tests are a valid surrogate for the more time-consuming professional evaluation

traditionally associated with outcome studies, but that obtaining patient-reported outcome data has significant implications for the surgical office. If PROMs test results are to be collected, the patient needs to be able to complete the questionnaires easily, and the privacy of the patient data must be ensured. As well, a change in the culture of the surgical office may be needed to incorporate PROMs and ensure the information collected is made available to the surgeon.

Rolfson and colleagues¹⁴ found that patients asked to complete a pen-and-paper PROMs assessment were twice as likely to respond as patients asked to complete a questionnaire electronically using the Internet. Similarly, Beach and colleagues⁶ outlined the need for data to be collected in a regular and easy manner and the importance of privacy protection in the use of data.

Incorporating the PROMs assessment into the patient intake process and ensuring that any tests are completed following treatment will mean a time commitment for the surgeon's medical office assistant (MOA) and may require a change in the office culture and the attitudes of the surgeon and MOA. Similarly, ensuring that the information collected is made

available to the surgeon will require time and effort. Initially, a PROMs score should be available at the time of the patient examination. Later, the cumulative scores for the surgeon's operations assessed by PROMs tests should be reported back to the surgeon every 6 months, in the same way that the surgeon's individual infection rate is reported by the health authority. The results should also be reported in an anonymous fashion to the bodies responsible for funding the treatment as a way to ensure that the patient is satisfied with the quality and necessity of the treatment and that funding the treatment is appropriate.

Recommendations

Baseline patient-reported outcome data should be obtained electronically and automatically as an integral part of patient intake. Patient self-assessment is a valid indicator of outcome and should be a regular part of evaluation. Corroboration by the use of a structured clinical evaluation would then only be necessary where validation of the PROMs instrument is needed and only for as long as it is necessary to validate the instrument in question.

The value of incorporating PROMs material into clinical practice is well

Instruments used in study of patient-reported outcome measures

Subjective tests

EuroQol 5D-SL: EQ-5D
EQ-5D visual analog scale: EQ-VAS
Oxford hip score: OHS
Oxford knee score: OKS
Oxford shoulder score: OSS

Objective tests

Constant shoulder score: CSS
Harris hip score: HSS
Knee Society score: KSS

outlined by Ayers and colleagues,¹⁵ who state that "Integrating PROs into routine orthopaedic patient visits can provide key information to monitor changes in symptom severity over time, support shared clinical care decisions, and assess treatment effectiveness for quality initiatives and value-based reimbursement." The same would hold true for all medical interventions, surgical and nonsurgical.

Support should be provided for instituting patient-reported outcome measurement in BC.

The province should take a leading role in implementing the use of various PROMs instruments, both general and specific types, in the offices of surgeons and other practitioners. Implementation would involve helping with statistical analysis and establishing ways to obtain consent and handle collection and sharing of patient data to ensure privacy. The data obtained regarding surgical outcomes would help the Ministry of Health with ongoing oversight and provide information about the effectiveness of procedures. The benefits would offset the cost of supporting the use of PROMs in surgeon's offices and contribute to a dramatic change in the way surgical care is provided in BC. It would be

The results should also be reported ... to the bodies responsible for funding the treatment ... to ensure that the patient is satisfied with the quality and necessity of the treatment.

reasonable to offer all surgical specialties support for implementing a PROMs project.

Competing interests

None declared.

References

1. McGrail K, Bryan S, Davis J. Let's all go to the PROM: The case for routine patient-reported outcome measurement in Canadian healthcare. *Healthc Pap* 2012;11:8-18.
2. Kalairajah Y, Azurza K, Hulme C, et al. Health outcome measures in the evaluation of total hip arthroplasties—a comparison between the Harris hip score and the Oxford hip score. *J Arthroplasty* 2005; 20:1037-1041.
3. Maher AJ, Kilmartin TE. Patient reported outcomes following the combined rotation scarf and Akin's osteotomies in 71 consecutive cases. *Foot (Edinb)* 2011; 21:37-44.
4. Jenkins PJ, Clifton R, Gillespie GN, et al. Strength and function recovery after multiple-ligament reconstruction of the knee. *Injury* 2011;42:1426-1429.
5. Harwin SF, Issa K, Given K, et al. Clinical and patient-reported outcomes of primary TKA with a single-radius design. *Orthopedics* 2013;36:e877-882.
6. Beach WR, Browne JP, Bastaki H, et al. Collecting outcomes data in the private practice setting: Why is it important and how to get the data without disturbing the workflow. *Sports Med Arthrosc* 2013; 21:148-151.
7. Browne JP, Bastaki H, Dawson J. What is the optimal time point to assess patient-reported recovery after hip and knee replacement? A systematic review and analysis of routinely reported outcome data from the English patient-reported outcome measures programme. *Health Qual Life Outcomes* 2013;11:128.
8. Judge A, Arden NK, Kiran A, et al. Interpretation of patient-reported outcomes for hip and knee replacement surgery: Identification of thresholds associated with satisfaction with surgery. *J Bone Joint Surg Br* 2012;94:412-418.
9. Parker SL, Adogwa O, Paul AR, et al. Utility of minimum clinically important difference in assessing pain, disability, and health state after transforaminal lumbar spondylolisthesis. *J Neurosurg Spine* 2011;14:598-604.
10. Wright CJ, Chambers GK, Robens-Paradise Y. Evaluation of indications for and outcomes of elective surgery. *CMAJ* 2002;167:461-466.
11. de Boer TA, Gietelink DA, Vierhout ME. Discrepancies between physician interview and a patient self-assessment questionnaire after surgery for pelvic organ prolapse. *Int Urogynecol J Pelvic Floor Dysfunct* 2008;19:1349-1352.
12. Harris IA, Harris AM, Naylor JM, et al. Discordance between patient and surgeon satisfaction after total joint arthroplasty. *J Arthroplasty* 2013;28:722-727.
13. Bream E, Charman SC, Clift B, et al. Relationship between patients' and clinicians' assessments of health status before and after knee arthroplasty. *Qual Saf Health Care* 2010;19:e6.
14. Rolfson O, Salomonsson R, Dahlberg LE, et al. Internet-based follow-up questionnaire for measuring patient-reported outcome after total hip replacement surgery—reliability and response rate. *Value Health* 2011;14:316-321.
15. Ayers DC, Zheng H, Franklin PD. Integrating patient-reported outcomes into orthopaedic clinical practice: Proof of concept from FORCE-TJR. *Clin Orthop Relat Res* 2013;471:3419-3425. **BCMJ**

The use of PROMs [could] contribute to a dramatic change in the way surgical care is provided in BC.