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

Michael Stanger, MD, Christine Morrison, PT, Erdem Yazganoglu, MD, Munjeet (Margi) Bhalla, PhD, MPA

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.

This article has been peer reviewed.
A study to assess patient-reported outcome measures (PROMs) and to investigate the practicality of using PROMs in a surgical office

Background
McGrail and colleagues 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, few of these studies have originated in a non-academic setting.

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 information, 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. 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.

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

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 EQ-5D and Oxford hip score values before (a) and after (b) surgery.](image1)

![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.](image2)
A study to assess patient-reported outcome measures (PROMs) and to investigate the practicality of using PROMs in a surgical office

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.

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

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

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

**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 postoperatively and at 12 months postoperatively. For hip patients, Judge and colleagues 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 colleagues 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 questionnaire: “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.” An acceptable number of patients reaching an MCID would “imply clinical significance and justification for implementation into clinical practice.”

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 comparing 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.

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

<table>
<thead>
<tr>
<th>Test</th>
<th>Pre-op values</th>
<th>Post-op values at 3–6 months</th>
<th>Change from pre-op to post-op values at 3–6 months</th>
<th>Post-op values at 12 months</th>
<th>Change from pre-op to post-op values at 12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ-SD (max 1.0)</td>
<td>0.43</td>
<td>0.74</td>
<td>+0.31</td>
<td>0.73</td>
<td>+0.30</td>
</tr>
<tr>
<td>EQ-VAS (max 100)</td>
<td>65.04</td>
<td>79.75</td>
<td>+14.71</td>
<td>78.88</td>
<td>+13.84</td>
</tr>
<tr>
<td>OHS (max 48)</td>
<td>20.32</td>
<td>37.20</td>
<td>+16.88</td>
<td>41.55</td>
<td>+21.23</td>
</tr>
<tr>
<td>HSS (max 100)</td>
<td>47.15</td>
<td>82.00</td>
<td>+34.85</td>
<td>89.66</td>
<td>+42.51</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Test</th>
<th>Pre-op values</th>
<th>Post-op values at 3–6 months</th>
<th>Change from pre-op to post-op values at 3–6 months</th>
<th>Post-op values at 12 months</th>
<th>Change from pre-op to post-op values at 12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ-SD (max 1.0)</td>
<td>0.53</td>
<td>0.73</td>
<td>+0.20</td>
<td>0.68</td>
<td>+0.15</td>
</tr>
<tr>
<td>EQ-VAS (max 100)</td>
<td>70.27</td>
<td>78.00</td>
<td>+7.73</td>
<td>77.53</td>
<td>+7.25</td>
</tr>
<tr>
<td>OKS (max 48)</td>
<td>22.87</td>
<td>34.10</td>
<td>+11.23</td>
<td>37.72</td>
<td>+14.85</td>
</tr>
<tr>
<td>KSS (max 100)</td>
<td>47.13</td>
<td>82.15</td>
<td>+35.02</td>
<td>90.58</td>
<td>+43.45</td>
</tr>
</tbody>
</table>

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

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 colleagues\(^{10}\) 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 colleagues\(^{11}\) 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 colleagues\(^{12}\) found a discrepancy of only 4% between the surgeon-reported outcome and the patient-reported assessment. Bream and colleagues\(^{13}\) also found a reasonable correlation between the physician’s evaluation and the patient’s.

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 number 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

<table>
<thead>
<tr>
<th>Measure</th>
<th>Hip replacement</th>
<th>Knee replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary</td>
<td>Revision</td>
</tr>
<tr>
<td>Joint replacements at or above MCID*</td>
<td>177</td>
<td>3</td>
</tr>
<tr>
<td>Assessments where difference can be calculated</td>
<td>287</td>
<td>16</td>
</tr>
<tr>
<td>Percentage of joint replacements at or above MCID*</td>
<td>61.7%</td>
<td>18.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>Hip replacement</th>
<th>Knee replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary</td>
<td>Revision</td>
</tr>
<tr>
<td>Joint replacements at or above MCID*</td>
<td>120</td>
<td>3</td>
</tr>
<tr>
<td>Assessments where difference can be calculated</td>
<td>149</td>
<td>11</td>
</tr>
<tr>
<td>Percentage of joint replacements at or above MCID*</td>
<td>80.5%*</td>
<td>27.3%</td>
</tr>
</tbody>
</table>

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

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

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

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 colleagues\textsuperscript{14} 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 colleagues\textsuperscript{6} 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 outlined by Ayers and colleagues\textsuperscript{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

### 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

---

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

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

Competing interests
None declared.

References
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.