

# RESEARCH METHODS & REPORTING

## Routine use of patient reported outcome measures in healthcare settings

Jill Dawson,<sup>1,2</sup> Helen Doll,<sup>1</sup> Ray Fitzpatrick,<sup>1</sup> Crispin Jenkinson,<sup>1</sup> Andrew J Carr<sup>3</sup>

The use of patient reported outcome measures might seem to be quite straightforward, but several pitfalls await clinicians with limited expertise. **Jill Dawson and colleagues** provide a guide for individuals keen to use patient reported outcome measures locally

Patient reported outcome measures (PROMs) are standardised, validated questionnaires that are completed by patients to measure their perceptions of their own functional status and wellbeing. Many such measures were originally designed for assessing treatment effectiveness in the context of clinical trials,<sup>1</sup> but are now used more widely to assess patient perspectives of care outcomes. This outcomes based definition of PROMs distinguishes them from questionnaires used to measure patients' experience of the care process.

PROMs are designed to measure either patients' perceptions of their general health ("generic" health status) or their perceptions of their health in relation to specific diseases or conditions. The short form 36 (SF-36) health survey,<sup>2</sup> for example, is a generic questionnaire that assesses self perceived health status by using 36 questions relating to eight broad areas (or "domains") of wellbeing. Examples of condition specific questionnaires include the Parkinson's disease questionnaire (PDQ-39),<sup>3</sup> which assesses quality of life in patients with Parkinson's disease; the visual function questionnaire (VF-14),<sup>4</sup> which uses 14 questions to measure various aspects of visual function affected by cataracts; and

the Oxford hip score,<sup>5</sup> which uses 12 questions to assess hip pain and function in relation to outcomes of hip replacement surgery.

Patients complete PROMs by rating their health in response to individual questions. These responses are scored (from 0 to 4, for example) according to the level of difficulty or severity reported by the patient. When PROMs are analysed, the individual ratings are combined to produce an overall score to represent an underlying phenomenon or "construct," such as "perceived level of pain" or anxiety. The analysis of PROMs tends to focus on the amount of change that has occurred in the patients' condition or their general health related quality of life, as represented by a change in PROM score following an intervention.

To date, PROMs have been used in clinical trials,<sup>6,7</sup> national audits,<sup>8</sup> and registers for joint replacement<sup>9,10</sup> and other conditions.<sup>11</sup> However, the routine use of PROMs has become widespread in health care at a local level.<sup>12</sup> Interest is also rapidly growing in the application of PROMs in the context of audit and "registers," to inform individual care and manage the performance of healthcare providers.<sup>12-16</sup> Indeed, in the specific areas of hip and knee replacement, inguinal hernia repair, and varicose vein surgery, the routine collection of PROMs has, since April 2009, been introduced throughout the NHS to measure and improve clinical quality.<sup>17</sup> Government led initiatives such as this are likely to encourage more widespread use of PROMs at both a national and a local level.

Specific guidance on methods for collecting baseline PROM data are provided in guidelines for the recent NHS-wide PROMs initiative,<sup>18</sup> in which subsequent data collection and handling are undertaken by private contractors. This article, however, is aimed at individuals who are keen to use PROMs at a local level, who may have limited research experience or access to expertise and advice on relevant research methods, and who may be unaware of a number of pitfalls that could undermine their aim of ultimately producing useful, meaningful data. In addition, there are very few published examples of the application of PROMs in the context of clinical governance and quality assurance,<sup>19</sup> with this form of application being largely unevaluated. Evidence of the impact of using PROMs on routine practice is also lacking.

### Using an appropriate validated measure

When choosing a PROM to use, careful consideration should be given to the content of the questionnaire and its relevance

<sup>1</sup>Department of Public Health, University of Oxford, Oxford OX3 7LF

<sup>2</sup>Oxford Brookes University, School of Health and Social Care, Oxford OX3 0BP

<sup>3</sup>Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford OX3 7LD

Correspondence to: J Dawson  
jill.dawson@dphpc.ox.ac.uk

Accepted: 7 September 2009

Cite this as: *BMJ* 2010;340:c186  
doi: 10.1136/bmj.c186

### SUMMARY POINTS

Patient reported outcome measures (PROMs) are standardised, validated questionnaires that are completed by patients to measure their perceptions of their own functional status and wellbeing

An appropriate and validated measure that is suitable for both the particular study population and the reason for collecting the PROMs data should be chosen

PROMs data need to be obtained from relevant patients at the same point in time relative to the date of an intervention or event of interest (for example, within four weeks pre-intervention, then at six months following the intervention) and recorded in association with the date of completion (not the date of data entry)

The intensity with which follow-up information is sought and obtained is known to greatly influence study results; every effort should thus be made to minimise missing data and the biases that might otherwise occur

Poor data cannot be "fixed" in an analysis by a statistician. Advice should be sought from those with relevant expertise from the very beginning of the study



ILLUSTRATION: JILL DAWSON, ADDITIONAL TECHNICAL ASSISTANCE: PHILIP SAUNDERS

to the intended form of usage and patient group. An appropriate measure is one that is supported by published evidence demonstrating that it is acceptable to patients, reliable, valid, and responsive (sensitive to change).<sup>1</sup> In addition, evidence for these properties needs to have been obtained in a similar context and on similar types of patients (in terms of age range, sex, and diagnostic or surgical category) to those whom the PROM is now to be applied. Using a PROM that meets these criteria is likely to maximise the response rate.

Choosing the right PROM for a particular purpose can be challenging because there may be a number of relevant questionnaires from which to choose. Alternatively, none may seem entirely appropriate as potential measures may include a number of questions irrelevant to the study sample—questions about sports participation or vigorous physical activity, for example, may not suit most elderly people. Listings of available measures<sup>20</sup> and systematic reviews of available instruments can assist in selecting an appropriate PROM.

Once a seemingly appropriate instrument has been identified, it is advisable to pilot the questionnaire on a small

number of patients. This process can reveal whether or not the questionnaire truly is appropriate for the intended purpose. For instance, a questionnaire will be unsuitable if the questions address the patients' state of mind "today" and patients are likely to complete the questionnaire on the day that they are admitted for treatment—a time when they may be unusually anxious.

It is important to note that the wording of a validated PROM should not be changed because even relatively small alterations can make a considerable difference to the meaning of the questions and consequently to the measurement properties of a questionnaire.

### Data collection and storage

PROMs are generally applied in longitudinal studies that have at least one follow-up survey planned. Good research practice requires investigators to clearly identify the purpose of the study, and data should be collected at prespecified time points so that particular questions (for example, how successful a procedure is at one year after a particular intervention) can be addressed. In the absence of a precise research question (for example, in exploratory research or descriptive audits), a reason for collecting PROMs data, preferably in relation to an event (for example, a particular intervention with a date), and any follow-up period still need to be specified before commencing data collection. This approach will help guide and standardise methods of data collection and aid the design of any associated database for storing data, as well as inform consideration of inclusion and exclusion criteria. If PROMs are collected to monitor long term conditions (for example, diabetes) where there is no specific "event," or in situations where there is no prospect of obtaining both pre-intervention and post-intervention assessments (for example, shortly after a stroke), a different rationale for the timing of regular assessments is required.<sup>21</sup>

Plans for long term data collection may naturally lead to other considerations for data gathering and storage. For instance, conditions and interventions that can affect bilateral structures (such as joints, eyes, or breasts), or that may require subsequent therapy revision or more than one course of treatment, can create complexity at every stage of data collection and storage and, indeed, when commencing analysis. The unit of analysis (that is, patient *v* right or left joint, eye, or breast) should preferably be decided upon in advance and any database designed accordingly.

Dates are crucial to longitudinal outcomes analysis, but they need to be the right ones. PROM questionnaires need to be obtained and responses recorded with the date of completion—not the date of data entry, which may involve a time lag—and with reference (labelled with and/or linked) to the date of an intervention or event of interest (for example, date of surgery, admission for rehabilitation, or start (or end) of a course of chemotherapy). Staff conducting data entry will need to be trained in relation to the importance of these issues.

Methods of data collection should be piloted and reviewed at an early stage. Once practicable methods have been tried and tested, they should be written down and adhered to. All these steps, as well as detailing methods for informing patients about the project and obtaining their written consent to participate, will be necessary if the approval of



ILLUSTRATION: JILL DAWSON, ADDITIONAL TECHNICAL ASSISTANCE: PHILIP SAUNDERS

**Box 1 | Is the sample representative?**

A response rate of 80% at baseline sounds very acceptable, particularly if the response to the first follow-up survey is also 80%. If the non-responders at each stage are different people, however, these values would equate to only a 60% overall response rate for measures of change (which require the presence of both pre-intervention and post-intervention measures of outcome). This rate would not be considered adequate in terms of sample representativeness.

an institutional or external research ethics committee is required.

PROMs are meant to represent the patients' perspective and be independent of the views of the clinical team providing their care. The method of data collection should, therefore, ensure that patients self complete their questionnaire unobserved and unaided by members of the clinical team. Assistance with questionnaire completion from a relative or friend, however, is occasionally unavoidable and indeed helpful. Nevertheless, a patient's inability to understand a questionnaire, for reasons of impaired cognition or difficulty with the language in which it is available, should constitute an exclusion criterion.

Translation of PROMs into other languages involves establishing conceptual and semantic equivalence, a task not to be undertaken lightly. This process should include forward and backward translation methods, plus an assessment of the translated questionnaire's measurement properties. The accepted method of translating and re-evaluating a PROM is both demanding and costly, so most PROMs are not available in a variety of different languages. This can prove to be problematic in healthcare settings that serve populations with diverse language preferences. Asking a relative or friend to translate the questionnaire for the patient is not acceptable, as a faithful translation that maintains the correct meaning cannot be guaranteed.

Data should be stored in a database or spreadsheet in a manner that allows for immediate statistical analysis without the need for detective work and complex data programming—that is, stored in an unambiguous fashion and with variables appropriately labelled. The aim should be to minimise complexity—for example, by avoiding the use of relational databases, which can add additional complexity to an already complicated process. In addition, methods for downloading data and conducting some simple analyses should be piloted before too many cases (no more than, say, 20) have been entered.

**Minimising missing and duplicated data**

The most successful trials that use PROMs are undoubtedly those that achieve very high questionnaire response rates at the prespecified times.<sup>22</sup> Nevertheless, systems to maximise the number of questionnaire returns carry cost implications,<sup>22 23</sup> and a balance has to be struck between maximising response and alienating patients.

Responders may differ systematically from non-responders in ways that matter—for example, they may have poorer general health or represent a particular age band or socioeconomic group.<sup>24 25</sup> Thus every effort should be made to address such potential biases. Where PROM data are to be

obtained by post, sending patients a reminder letter if questionnaires are not returned within two or three weeks (with a contact telephone number in case patients need to request a second questionnaire) is generally essential to obtain satisfactory response rates from a representative sample of the population (box 1).

Collecting follow-up data when patients attend outpatient appointments is inadvisable because of the risk of introducing bias. Outpatient appointments can rarely be organised to occur at precise time points after a hospital based procedure or course of treatment, and are frequently changed by the hospital or the patient. Also, patients who experience continuing problems are more likely to attend, or attend more often, than other patients, which could mean extra data are obtained from patients with poorer outcomes. It is in any case much easier to regularise and monitor the collection of follow-up data if questionnaires are sent out to patients' homes from one office on relevant dates, with the dates when questionnaires were sent out and returned then recorded in a database.

Follow-up times should be the same for all patients in relation to the intervention or other key event. Collecting data continuously but irregularly after an intervention (that is, not at particular time points) will seriously limit the usefulness of the data (for an example, see Saleh et al<sup>23</sup>). This can easily happen if follow-up data are collected when patients attend outpatient appointments.

**Thinking about data analysis**

Before commencing data collection, serious consideration should be given to the way in which data will ultimately be analysed. This process will help to identify other pieces of information that may need to be collected to place the PROMs data in an appropriate context and to interpret the data correctly. For instance, outcomes might be expected to suggest that an intervention is less successful for some patients than for others—for example, hip replacement may not fully restore a patient's mobility if the patient has another coexisting condition that affects walking ability. In this example, details about other conditions that might affect walking must be obtained during follow-up to allow adjustment to be made for such factors in the analysis, in addition to collecting outcomes data specific to the hip operation (box 2).

The importance of obtaining additional information from patients needs to be weighed carefully against the risk of missing data owing to patients feeling overburdened by a lengthy questionnaire and not completing it fully.

If data collection has occurred over a number of years, a large amount of data will be available. It is important to

**Box 2 | What is the influence of case mix on PROMs?**

The analysis and interpretation of results from PROMs used in an audit or study with a non-randomised design is complex because it is difficult to control for all the possible "case mix" factors that may influence outcomes. Some examples are presence of other comorbidities, severity of the condition before treatment commenced, period of time since start (or end) of treatment, between-subject variation in treatment (such as drug dosages), and previous or concurrent other forms of treatment.

**bmj.com: recent Research Methods & Reporting articles**

- ▶ Meta-analysis of individual participant data: rationale, conduct, and reporting (2010;340:c221)
- ▶ The routine use of patient reported outcome measures in healthcare settings (2010;340:c186)
- ▶ Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers (2010;340:c181)
- ▶ Good publication practice for communicating company sponsored medical research: the GPP2 guidelines (2009;339:b4330)
- ▶ Use of multiperspective qualitative interviews to understand patients' and carers' beliefs, experiences, and needs (2009;339:b4122)



recognise, however, that a large amount of data does not necessarily equate with good data. Poor (that is, biased) data cannot be “fixed” in an analysis, even by the cleverest of statisticians. Indeed, leading geneticist and statistician R A Fisher (1890-1962) once said: “To consult the statistician after an experiment is finished is often merely to ask him to conduct a postmortem examination. He can perhaps say what the experiment died of.”<sup>26</sup> We would, therefore, advocate seeking advice from those with relevant expertise from the beginning of the data collection period.

### Conclusions

Overall, many clinicians are very positive about the usefulness of collecting PROMs; this consensus is reflected in the widespread use of such measures. PROMs can be used to assess the impact healthcare interventions have on patients, assist with guiding resource allocation, evaluate the effects of changes to services, and provide feedback to consultants to assist clinical governance. The systematic use of PROMs may result in improvements to patient outcomes in a number of ways—for example, by providing patient centred information and thus facilitating improved communication between doctors and their patients. Patients may also feel that healthcare personnel are more involved in their care because professionals are showing an interest in obtaining their perspective on their health and wellbeing.

The analysis of PROMs data may also reveal important differences in outcomes between different patient groups, which can trigger a subsequent more focused investigation. PROMs that are routinely collected are unlikely to reliably reveal the reasons underlying any such differences, however, given the difficulty of adjusting for all relevant confounders. In addition, it is important to be aware of the limitations of this new approach in influencing health care. The incautious application of PROMs may produce meaningless or misleading and potentially harmful results. Many of the points raised in this paper represent pitfalls that are easy to fall into, but that are also largely avoidable if sufficient time and thought occur at the planning stage.

We acknowledge the additional technical assistance with illustrations that was provided by Phillip Saunders, Unit Administrator, Department of Public Health (Health Services Research Unit), University of Oxford.

**Contributors:** All the authors have considerable experience in developing, evaluating, and applying questionnaires for patients and are currently involved in long term multicentre trials where patient reported outcomes are the main end points. JD and AJC have chiefly worked in the area of patient reported outcomes in the context of orthopaedic surgery. HD is a senior statistician specialising in the development and application of patient reported outcome measures and on randomised controlled trials of complex interventions. RF has worked on both patient reported outcomes and patient experience of care relating to a wide range of conditions and interventions, both surgical and long term medical. CJ has worked on both patient reported outcomes and patient experience of care, the latter related to his work with the Picker Institute Europe, Oxford, UK. All authors contributed to the writing of this paper. JD is the guarantor.

**Competing interests:** All authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare (1) No financial support for the submitted work from anyone other than their employer; (2) No financial relationships with commercial entities that might have an interest in the submitted work; (3) No spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; (4) No non-financial interests that may be relevant to the submitted work.

**Provenance and peer review:** Commissioned, externally peer reviewed.

- Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2:1-74.
- Ware-JE J, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
- Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. The Parkinson's Disease Questionnaire (PDQ-39): development and validation of a Parkinson's disease summary index score. *Age Ageing* 1997;26:353-7.
- Steinberg EP, Tielsch JM, Schein OD, Javitt JC, Sharkey P, Cassard SD, et al. The VF-14: An index of functional impairment in patients with cataract. *Arch Ophthalmol* 1994;112:630-8.
- Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg [Br]* 1996;78:185-90.
- Silverman LR, Demakos EP, Peterson BL, Kornblith AB, Holland JC, Odchimar-Reissig R, et al. Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the Cancer and Leukemia Group B. *J Clin Oncol* 2002;20:2429-40.
- Grant AM, Wileman SM, Ramsay CR, Mowat NA, Krukowski ZH, Heading RC, et al for the REFLUX Trial Group. Minimal access surgery compared with medical management for chronic gastro-oesophageal reflux disease: UK collaborative randomised trial. *BMJ* 2008;337:a2664.
- Williams O, Fitzpatrick R, Hajat S, Reeves BC, Stimpson A, Morris R, et al. Mortality, morbidity, and 1-year outcomes of primary elective total hip arthroplasty. *J Arthroplasty* 2002;17:165-71.
- Malchau H, Garellick G, Eisler T, Herberts P. Presidential guest address: the Swedish Hip Registry: Increasing the sensitivity by patient outcome data. *Clin Orthop Rel Res* 2005.
- New Zealand Joint Registry. *The New Zealand Joint Registry Nine Year Report January 1999 to December 2007*. Department of Orthopaedic Surgery and Musculoskeletal Medicine, Christchurch Hospital, 2008. [www.cdhb.govt.nz/njr/](http://www.cdhb.govt.nz/njr/).
- Zanoli G, Nilsson LT, Stromqvist B. Reliability of the prospective data collection protocol of the Swedish Spine Register: test-retest analysis of 119 patients. *Acta Orthop* 2006;77:662-9.
- Appleby J, Devlin N. *Measuring success in the NHS. Using patient-assessed health outcomes to manage the performance of healthcare providers*. The King's Fund, 2004.
- Wasson J, Keller J, Rubenstein L, Hays R, Nelson E, Johnson D. Benefits and obstacles of health status assessment in ambulatory settings: the clinician's point of view. *Med Care* 1992;30:MS42-MS49.
- Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? *Soc Sci Med* 2005;60:833-43.
- Haywood K, Marshall S, Fitzpatrick R. Patient participation in the consultation process: a structured review of intervention strategies. *Patient Educ Couns* 2006;63:12-23.
- Timmins N. Assessing patient care—NHS goes to the PROMS. *BMJ* 2008;336:1464-5.
- Department of Health. *Our NHS, our future: NHS next stage review. Interim report*. 2007. [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_079077](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_079077).
- Department of Health. *Guidance on the routine collection of Patient Reported Outcome Measures (PROMs)*. 2009. [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_092647](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_092647).
- Vallance-Owen A, Cubbin S, Warren V, Matthews B. Outcome monitoring to facilitate clinical governance: experience from a national programme in the independent sector. *J Public Health* 2004;26:187-92.
- MAPI Research Trust. *Patient reported outcome and quality of life instruments database*. 2009. <http://www.proqolid.org>.
- Black N, Jenkinson C. Measuring patients' experiences and outcomes. *BMJ* 2009;339:b2495.
- Ganz PA, Gotay CG. Use of patient-reported outcomes in phase III cancer treatment trials: lessons learned and future directions. *J Clin Oncol* 2007;25:5063-9.
- Saleh KJ, Bershady B, Cheng E, Kane R. Lessons learned from the hip and knee musculoskeletal outcomes data evaluation and management system. *Clin Orthop Rel Res* 2004;429:272-8.
- Bracken M. Reporting observational studies. *Br J Obstet Gynaecol* 1989;96:383-8.
- Matthews FE, Chatfield M, Freeman C, McCracken C, Brayne C. Attrition and bias in the MRC cognitive function and ageing study: an epidemiological investigation. *BMC Public Health* 2004;4:12.
- Edwards AWF. Some quotations from R A Fisher. <http://www.economics.soton.ac.uk/staff/aldrich/fisherguide/quotations.htm>.