PROMs - How was it for you?

CR Howie and DF Hamilton (Edinburgh)

During the last decade patient reported outcome measures (PROMs) and satisfaction have become accepted as valid tools in orthopaedic research and healthcare policy. In 2008 Lord Darzi¹, in his landmark report on the NHS, set outcomes and quality at the centre of government policy, stating that “as important is the effectiveness of care from the patient’s own perspective which will be measured through PROMs”. Measuring treatment results is not a new concept: well over a century ago Florence Nightingale registered the outcomes of care as “Relieved; Unrelieved; and Died”. In 1908 E A Codman (an orthopaedic shoulder and tumour surgeon in Boston) asked, “What actually happened to the patient?”.

Codman’s ‘End result’ idea:

“The common sense notion is that every hospital should follow every patient it treats, long enough to determine whether or not the treatment has been successful, and then to inquire, “If not, why not?” with a view to preventing similar failures in the future.”

This question is equally pertinent in today’s healthcare environment, though how we determine success is important. The Oxford English Dictionary defines an outcome as a visible or practical result. Originally, our criteria for assessing surgical success were based on lack of complications (notably death) and achieving the surgical result we were aiming for (union, stability, osseo-integration etc.). While clearly important, death is a rare orthopaedic event; variation in death rates likely represents case selection, anaesthesia and aftercare, and almost certainly reflects hospital support services for our frail elderly fragility fracture patients. Perhaps this metric is representative of our attention to management issues in hospital rather than our orthopaedic abilities? Death rates cannot discern differences in clinical outcome following planned procedures which are carried out to address morbidity, and is therefore neither specific nor sensitive for the majority of planned orthopaedic procedures. Lord Darzi, in his evidence to the parliamentary health committee in 2008, highlighted that mortality statistics were an imprecise and perhaps inappropriate statistic to measure success in many conditions.

Surgical complications (infection, dislocation, non-union), while important and worth monitoring, are now low in orthopaedics, thanks to effective patient safety programmes. If such complications are rare (or present late), unfeasibly large numbers are needed in controlled trials to show significance. It can be difficult to collect all complication episodes or indeed agree that a complication has occurred: for example, infection, superficial or deep, depends on the quality of bacteriological diagnosis and Pulmonary Embolism depends on the vigilance of staff and the availability of Computer Tomography Pulmonary Angiography. To further confuse the issue, the Swedish arthroplasty register first noted that around half the patients with an infected joint replacement (a bad outcome in the surgeon’s eyes) were satisfied with the results of surgery. This is difficult to explain if we assume that lack of complication is the arbiter of success.

For arthroplasty, survival curves have rightly become the standard evaluation for implant longevity, all cause revision is generally accepted as the end-point yet unfortunately not universally quoted. This method of analysis has been criticised as neither adequately defining implant failure nor commenting on patient function prior to that failure². Equally, though apparently a firm and definite endpoint, the threshold for revision may vary from surgeon to surgeon. In addition case mix at primary surgery will affect the figures, and selective reporting (is a liner exchange a revision?) will also skew results. Finally, failure to revision is rare and often takes place after many years, thus (in the context of new technologies) problems of new failure modes are often not detected until a number of years later.

Assessment of ‘results’ has gradually shifted away from a simplistic success vs. failure model towards the use of quantifiable measures. Clinical outcomes such as range of motion, the clinician’s report of pain relief, radiographic...
Much better. The government is now more sophisticated in measuring health status of the patient.

Errors. We took our unreliable measurements and collated many different values into a compound score based on logical, well meaning, but arbitrary assessments of the importance of each individual item contributing to the score.

More recently our focus has shifted towards PROMs to ascertain the patient's perception of success. A PROM is any measure reported by the patient (in a validated standardised format). Proponents of these PROMs suggest they provide a remarkably sophisticated measure of whether a patient feels better and how much better. The government now requires that providers of planned surgery of NHS patients gather this information. Various scores are now used to compare prosthesis or techniques, and in auditing departments and individual surgeons. Generic health status questionnaires such as the EQ-5D, have been widely used and validated to look at the economic benefits of treatments. Unfortunately, these generic scores are insensitive to changes in one joint but reflect the general change in health status of the patient.

For specific operations more specific questions need to be used. There are many validated joint or disease specific scores available of which the most commonly used in this country are the various Oxford scores, the Manchester/Oxford foot score, DASH, and the Constant Shoulder score. Widely used they also have limitations, most notably a ceiling effect when a patient begins with a good score. The specificity of these questions is also debated, and again a number of extended scores are available for particular situations. Joint or disease specific PROMs are commonly used to assess ‘functional (or clinical) outcome’ as they are cheap, effective at collecting large volumes of data, and don’t require follow-up clinic visits as they can be returned by post or electronically. However, joint specific PROMs are heavily influenced by a patient’s pain levels. As opposed to being an assessment of functional activity, the patient reports their ‘experience’ of performing that activity. Despite these reservations, there is evidence that PROM scores can highlight potential implant failure.

If PROMs are used to measure outcome, what is it that influences this outcome (assuming all surgery and surgeons are equally successful)? Much has been written on this. Crudely, the worse the pre-operative scores, the worse the post-operative scores. Thus we need pre-and post-op data to assess this in a meaningful way. There are a number of factors known to predict outcome scores: the procedure and joint is one, hips and knees achieve different improvements (accepting the different scoring systems). Patient factors such as depression, obesity, age and co-morbidity (often co-existent in the one patient) are all identified in the literature as contributing to the patients’ reporting of pain and function. Social deprivation is also thought to contribute, though the exact combination and extent of these various factors are difficult to interpret. Interestingly, there is some evidence that we already take such factors into account in our patient selection (co-morbidities and obesity, for example) and perhaps the GP and ourselves wait longer before considering surgery in the ‘at risk’ patients as their scores are worse before we operate. When taking the factors known to affect change in scores into account there is surprisingly little difference around the country in the values achieved for joint replacement. Uncidentally, but unsurprisingly, given the cost and popularity of joint replacement, outcome scores have been used to limit access to healthcare ultimately to reduce cost. This will prove very difficult though as these scores were specifically developed as tools to measure treatment success, as opposed to diagnostic parameters of an individual’s suitability for surgery.

Satisfaction is a different metric. Though patient satisfaction is related to the clinical outcome score, it is not the same thing. Indeed some patients reporting a bad clinical outcome, in terms of pain and function, may report good levels of satisfaction with their surgical outcome and vice versa. We recently reviewed the factors associated with satisfaction in joint arthroplasty and found the only things that really mattered were meeting of pre-operative expectations, achieving pain relief and the hospital experience. Perhaps patient satisfaction is the most important outcome. This is well recognised in service industries, though remains something of a nebulous concept in clinical care. Because of recent, well publicised failings in the NHS, patient experience is the latest target for measurement. Patient Reported Experience Measures (PREMs) have been well validated but not widely used. They record the actual delivery of healthcare.

Drawing on the widespread use of surveys in the service industries to maintain quality services, the government plans to introduce the ‘friends and family test’ (Net Promoter Score as used in the service industries). Patients are asked whether or not they would recommend the care they have received to a loved one if they were in a similar situation. We have reviewed this measure in the context of joint arthroplasty and found that the same three factors determine this score: the meeting of expectations, achievement of pain relief (PROMs) and the experience of being in hospital (PREMs). Importantly, in terms of these satisfaction and experience measures, the patients’ demographics do not affect outcome in the same manner that they affect the PROMs, making them more suitable as an overarching metric. This generalisability is of course at the expense of specificity and detail. We know that patient satisfaction measures do not necessarily relate to the technical quality of the care received. Recent concern has been expressed that, if asked in isolation, these experience metrics provide little insight. We have found good correlation between satisfaction and ‘promoter score’ when asked in the context of a suite of questions about the episode of care. Assuming the absence of serious complication, a combination of PROMs, patient satisfaction and experience is perhaps the most appropriate way to determine ‘what actually happened to the patient’.

Orthopaedics has little to fear and everything to gain from using and publishing a balance of different PROMs to prove the value of our
care. Almost universally our results compare extremely favourably with other specialties and are actually better than many well respected service industries. We must remain focused on our patients and apply a little of what Lord Darzi described as “Starbucks style care”. The UK government has appointed Professor Don Berwick as Patient Safety ‘Czar’. He is a senior and well respected figure from the Boston Institute of Healthcare Improvement and latterly headed up the ‘Obamacare’ programme in the US. Berwick believes that E A Codman should be regarded as the father figure for measuring success in the healthcare environment. In the UK we can show that he is correct and that, in orthopaedics, we have learned from one of our own.

References

9 Hamilton DF, Lane JV, Gaston P, Patton JT, MacDonald D, Simpson AHRW, Howie CR. Is the ‘friends and family test’ (Net Promoter Score) a workable metric for the NHS and what does it measure? A PROSPECTIVE COHORT STUDY OF 6912 PATIENTS FOLLOWING TOTAL JOINT REPLACEMENT. [In Press].

Correspondence:
Mr Colin Howie
Dept. of Trauma and Orthopaedics
University of Edinburgh
Little France
Edinburgh EH16 4SU