YHEC

York Health Economics Consortium

What are PROMs

Adam B. Smith, PhD.
Project Director
York Health Economics Consortium
University of York

1. OVERVIEW

- **Patient-reported outcomes (PROs)** are evaluations made directly by patients concerning their health, symptoms, level of functioning, satisfaction with treatment, and so on without any input from their family, carers and healthcare professionals.

- **Patient-reported outcome measures (PROMs)** are the tools used to capture PRO information usually in the form of questionnaires. There are different types of PROMs, including **generic** instruments aimed at different patient populations as well as **disease-specific** and **treatment-specific** measures.

- **PROMs** are commonly used in clinical trials and are increasingly being used in clinical research and practice, for instance, to evaluate the impact and effect of novel drug treatments, monitor symptoms, and facilitate communication between patients and their healthcare professionals.

- Regulatory agencies such as the UK's National Institute for Health and Care Excellence (NICE) utilise **PROMs** for the cost-effectiveness analysis of new drugs and medical interventions.

- Bodies such as the Federal Drug Administration (FDA) in the United States and the European Medicines Agency (EMA) have set new benchmarks in terms of the properties required for **PROMs** when used to demonstrate patient benefit in clinical trials.

York Health Economics Consortium Ltd
Level 2 Market Square
University of York
Heslington
York YO10 5NH

T +44 (0)1904 323620
F +44 (0)1904 323628
E yhec@york.ac.uk
W www.yhec.co.uk

2. WHAT ARE PROMS

The last decade has seen an increasing focus and interest in capturing the patient experience in their interactions with (global) healthcare systems. Put simply **patient-reported outcomes (PRO)** are individuals' ratings of their quality of life, symptoms, treatment effects, functioning and so on, that are elicited directly from the patient without any input from their family, carers or healthcare professionals. The tools used to capture this information - usually questionnaires or surveys - are known as **patient-reported outcome measures (PROMs)**. To take some examples, PROMs are now virtually routinely included in randomized controlled trials to determine the impact of an intervention, e.g. a new drug or intervention (Basch et al., 2013). This is important to pharmaceutical companies as the information gained from PROMs may help to demonstrate the effectiveness of a drug. PROMs are also being incorporated into clinical practice to help monitor patients' symptoms and their experiences during treatment, as well as enhancing doctor-patient communication. For instance, studies have shown that the regular use of PROMs in oncology clinical practice leads to patients reporting symptoms more frequently to their physicians during long-term follow-up (Takeuchi et al., 2011). Since 2009 PROMs have been used to measure the effectiveness of surgical intervention in four indications in the UK's National Health Service (NHS). Patients' complete generic and disease-specific PROMs (see Section 3.1) prior to and after surgery for hip and knee replacement, as well as groin hernia and varicose vein repair (e.g. Gutacker et al., 2013).

3. TYPE OF PROMS

3.1 Generic, Condition and Treatment-Specific PROMs

There are different types of PROMs used to capture information from patients. The **general** or **generic** types of PROMs are not specific to any particular disease or condition and are intended primarily to allow comparisons across different patient groups. Two generic PROMs in widespread use are the SF-36 (Ware and Sherbourne, 1992) and the Nottingham Health Profile (Hunt et al., 1980) both of which comprise of questions about patients' psychological and physical health (e.g. emotional reactions/mental health, physical mobility/functioning), social functioning and pain. In contrast to this, PROMs such as the European Organisation for Research and Treatment of Cancer Quality-of-Life Questionnaire (EORTC QLQ-C30), the Kidney Disease Quality of Life-Long Form, and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) are examples of **disease-** or **condition-specific PROMs** to be used to assess symptoms, limitations etc., associated with medical conditions such as cancer, chronic kidney disease and osteoarthritis, respectively. For instance, the EORTC QLQ-C30 contains questions about the degree of pain, nausea and vomiting and fatigue that are specific and relevant to cancer patients. In addition to these general, condition-specific PROMs, there are also PROMs that are specific to treatment and symptoms associated with particular medical conditions. Examples of these **treatment- and symptom-specific PROMs** include the instruments from the FACIT Measurement System, such as the FACT-B and FACT-L which are specific to breast and lung cancer respectively, as well as the FAPSI-8 which is used to assess the symptoms associated with prostate cancer. Others, such as the FACT-BMT may be used for patients undergoing bone marrow transplantation.

3.2 Preference-Based Measures

The other main type of PROM is the **preference-based measure**, such as the EQ-5D and Health Utilities Index, which is sometimes also referred to as a **utility measure**. In contrast to the other types of PROM described above, patients' responses to questions on **preference-based PROMs** are not summed to provide a score, but rather they are used to form a profile score which is converted into an **index score** based on societal preferences for a given health state. Index scores from the EQ-5D may be used to calculate "Quality-Adjusted Life Years" (QALYs, see "What is a QALY?" for further information) which are utilized by bodies such as the UK's National Institute for Health and Care Excellence (NICE) in determining the cost-effectiveness of medical interventions. This and the scoring methods for other PROMs are described in more detail under **Response formats, scoring and domains** (see Section 4).

3.3 PREMs and PROM-PM

The final type of PROM now gaining more widespread use is the **patient-reported experience measure (PREM)**, which is often used to evaluate patients' experiences and assessment of the quality of care they have received. For instance, in the NHS PROMs programme referred to above, in addition to completing the EQ-5D, patients also answer additional post-operative questions about whether they had any allergic reactions to the drugs, urinary problems, bleeding and wound problems, and whether they needed to be re-admitted to hospital. Patients are also asked to rate the results of their operation (from "poor" to "excellent") and to assess the level of change in their health problems comparing pre- and post-surgery (from "much better" to "much worse"). A very recent development is to combine **patient-reported outcomes** with clinical **performance measures** to produce **patient-reported outcome performance measures (PRO-PM)** (Basch et al., 2014). **PRO-PMs** are an example of **where patient-reported outcomes**, for instance, pain or nausea and vomiting are being collected to help inform, monitor and evaluate healthcare services.

4. RESPONSE FORMATS, SCORING AND DOMAINS

PROMs are self-completed by patients with the most common response option being the **rating scale** (see Box 1). Rating scales normally consist of a descriptor reflecting, for instance, the degree to which patients rate their symptom level, such as pain. The descriptors correspond to numerical values along an **ordinal scale**, where there is a clear linear increase (or occasionally a decrease) in the degree of symptoms, quality of life, etc. Although there is a clear numerical difference between the **response categories** it is not correct to assume that the differences between response categories are equal, i.e. a rating of 4 - "very much" on a pain scale does not mean that the patient is experiencing twice as much pain as another who rated their pain as 2 – "a little". All that can be stated is that the first patient is experiencing more pain than the second. The number of response categories may differ both between as well as within individual PROMs. There is no restriction on the number of a response, although categories beyond 10 are rare. **Dichotomous** (2 response categories, i.e. yes/no) and 4 and 5 response categories (**polytomous response categories**) are the most common formats.

Box 1 shows two examples taken from the EORTC QLQ-C30. For the first question patients are asked to rate their level of pain experienced during the last week on a 4-point scale ranging from "Not at all" to "Very much". The second question asks patients to rate their overall quality of life on a 1 to 7 scale. For this question the descriptors have been removed and instead there are two **anchors** or reference points at the two extremes: "Very poor" and "Excellent".

Box 1. Example of response categories for the EORTC QLQ-C30

During the past week	**Not at All**	**A Little**	**Quite a Bit**	**Very Much**
Have you had pain?	1	2	3	4

How would you rate your overall quality of life during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

Although the scores on individual items may be used to indicate the patient's health status more often the scores are added together to produce an overall total (e.g. overall quality of life, psychological distress, etc.) or subgroups of items may be summed to produce scores for particular **domains**. For instance, the Hospital Anxiety and Depression Scale (HADS) was designed to assess the degree of psychological distress experienced by outpatients' attending hospital clinics. The 14 items of the HADS may be summed to form an overall psychological distress score. In addition to this, subgroups of the items may be summed separately to form separate scores for the anxiety and depression **domains**.

Another frequently used response format is the **visual analogue scale (VAS)** where patients are asked to indicate the severity or intensity of symptoms, for instance, by placing a mark along a horizontal or vertical line. An example of a VAS is shown in Box 2, although this particular scale also includes a numerical rating scale alongside the vertical VAS line (and two anchors) it is also possible for these to be presented without the numerical cues.

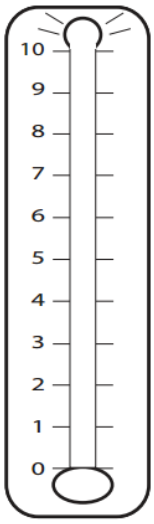
A final example of response category is shown in Box 3. This is taken from the Dartmouth COOP/WONCA instrument which was designed to be used in general practice, and where alongside the numerical rating scale there are also images to indicate, in this example, the degree of emotional problems experienced by patients. Scales such as this are commonly employed with paediatric patients, and may also be useful for patients with cognitive impairments and/or literacy problems.

Box 2. The Distress Thermometer

SCREENING TOOLS FOR MEASURING DISTRESS

Instructions: First please circle the number (0-10) that best describes how much distress you have been experiencing in the past week including today.

Extreme distress



No distress

Second, please indicate if any of the following has been a problem for you in the past week including today. Be sure to check YES or NO for each.

YES	NO	Practical Problems
❑	❑	Child care
❑	❑	Housing
❑	❑	Insurance/financial
❑	❑	Transportation
❑	❑	Work/school
❑	❑	Treatment decisions

YES	NO	Family Problems
❑	❑	Dealing with children
❑	❑	Dealing with partner
❑	❑	Ability to have children
❑	❑	Family health issues

YES	NO	Emotional Problems
❑	❑	Depression
❑	❑	Fears
❑	❑	Nervousness
❑	❑	Sadness
❑	❑	Worry
❑	❑	Loss of interest in usual activities

YES	NO	Spiritual/religious concerns
❑	❑	

YES	NO	Physical Problems
❑	❑	Appearance
❑	❑	Bathing/dressing
❑	❑	Breathing
❑	❑	Changes in urination
❑	❑	Constipation
❑	❑	Diarrhea
❑	❑	Eating
❑	❑	Fatigue
❑	❑	Feeling Swollen
❑	❑	Fevers
❑	❑	Getting around
❑	❑	Indigestion
❑	❑	Memory/concentration
❑	❑	Mouth sores
❑	❑	Nausea
❑	❑	Nose dry/congested
❑	❑	Pain
❑	❑	Sexual
❑	❑	Skin dry/itchy
❑	❑	Sleep
❑	❑	Substance abuse
❑	❑	Tingling in hands/feet

Other Problems: _____

_____

Box 3. Dartmouth COOP/WONCA Charts

FEELINGS

During the past 4 weeks...
How much have you been bothered by emotional problems such as feeling anxious, depressed, irritable or downhearted and blue?

Not at all		1
Slightly		2
Moderately		3
Quite a bit		4
Extremely		5

As noted above **preference-based measures** are scored differently to other **PROMs** producing a **profile** rather than a summary score. The EQ-5D-3L, for instance, consists of five domains, i.e. mobility, self-care, usual activities, pain/discomfort and anxiety/depression which are rated on a 1-3 scale. Patients' responses to the five domain result in a profile, e.g. 11111 indicates that the patient had no problems with any of the five domains, whereas 11223 would indicate no problems with mobility or self-care, some problems with usual activities and some pain/discomfort and severe anxiety/depression.

5. PSYCHOMETRIC PROPERTIES

As part of the development process PROMs are subjected to psychometric testing, in particular to determine whether the instruments are **valid, reliable** and **responsive** or **sensitive to change**. There are different types of **validity** and **reliability** the most important of which are **internal reliability (consistency), test-retest reliability and content validity**. **Internal reliability** assesses the consistency of items within the instrument, i.e. do they form a coherent whole? **Test-retest reliability** on the other hand measures whether the instrument – the PROM – performs consistently over time, i.e. does the instrument produce consistent evaluations of patients' outcomes? **Content validity** refers to whether the PROM in question is measuring the concepts, domains, symptoms and so on, that are relevant to patients. The majority of PROMs used in healthcare were developed as part of clinical research and may not have been intended for use in clinical trials and practice, and many, such as the EORTC QLQ-C30 were initially created without patient input. Recently guidance has been issued by regulatory authorities such as the Food and Drug Administration (FDA) in the USA and the European Medicines Agency (EMA) which have stressed the importance of involving patients in the early stages of developing new and evaluating existing PROMs for use in clinical trials. As a result of this guidance it is likely that a number of PROMs currently in use will not be deemed fit for purpose by the FDA and/or EMA for collecting PRO information to demonstrate patient benefit in clinical trials. Finally, another important psychometric property of PROMs is their **ability to detect change**, in other words whether they are **sensitive or responsive to change**. Often it is the change in scores on PROMs that is of importance rather than the actual score in helping to determine whether a medical intervention has been successful. For instance, in the NHS PROMs programme the change in score from baseline, i.e. before surgery, on the Oxford Knee Score is used alongside the clinical data to evaluate whether the intervention has led to improved knee function, mobility, reduction in pain, etc. from the patients' perspective. Linked to this is the **minimal important difference** or **MID**. There is no agreed definition of what constitutes an MID, however, most authors interpret it to mean the smallest change score perceived by patients as important (either beneficial or detrimental) and which would lead a patient or their doctor to consider a change in treatment or management.

6. CONCLUSION

PROMs are gaining increasing prominence in healthcare systems around the world particularly in terms of capturing the patient experience whether as part of clinical trials, clinical practice or research. There is increasing recognition from a regulatory and clinical perspective that patient input is critical to the management and treatment of patients to help inform clinical decision-making processes, as well as in the evaluation of novel medical interventions. In addition to these uses for PROMs, instruments are also being developed and employed, for instance, to enable patients to report adverse events in clinical trials (PRO - Common Terminology Criteria for Adverse Events). Existing measures may not fully comply with regulatory requirements, however, one area in which their use has been recommended is clinical comparative effectiveness research aimed at evaluating the relative benefits and harms between treatments aimed at the same indication.

After a number of decades of gradual evolution, the field of **PROMs** is currently experiencing renewed interest and rapid developments in terms regulatory and clinical guidance and recommendations, as well as technological advancements, such as **electronic PROMs** which are slowly being introduced (see Box 4).

Box 4. Electronic patient-reported outcome measures (ePROMs)

- **The benefits of collecting PRO information electronically (ePROMs) include:**
  - Rapid data acquisition
  - Less missing data
  - Easier and quicker storage of information
  - Facilitates data analysis
  - Reduced cost (no paper questionnaires, fewer staff required to collect completed forms)

- **However,**
  - The cost may be too high for smaller healthcare organisations and systems
  - Data security issues
  - Difficult to implement in a global clinical trial

7. REFERENCES

Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith ML, Coons SJ, Sloan J, Wenzel K, Chauhan C, Eppard W, Frank ES, Lipscomb J, Raymond SA, Spencer M, Tunis S. Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. J Clin Oncol. 2012; 30(34):4249-55.

Basch E, Autio K, Ryan CJ, Mulders P, Shore N, Kheoh T, Fizazi K, Logothetis CJ, Rathkopf D, Smith MR, Mainwaring PN, Hao Y, Griffin T, Li S, Meyers ML, Molina A, Cleeland C. Abiraterone acetate plus prednisone versus prednisone alone in chemotherapy-naive men with metastatic castration-resistant prostate cancer: patient-reported outcome results of a randomised phase 3 trial. Lancet Oncol. 2013; 14(12):1193-9.

Basch E, Snyder C, McNiff K, Brown R, Maddux S, Smith ML, Atkinson TM, Howell D, Chiang A, Wood W, Levitan N, Wu AW, Krzyzanowska M. Patient-reported outcome performance measures in oncology. J Oncol Pract. 2014; 10(3):209-11.

Gutacker N, Bojke C, Daidone S, Devlin N, Street A. Hospital variation in patient-reported outcomes at the level of EQ-5D dimensions: evidence from England. Med Decis Making. 2013; 33(6):804-18.

Holland JC, Andersen B, Breitbart WS, Compas B, Dudley MM, Fleishman S, Fulcher CD, Greenberg DB, Greiner CB, Handzo GF, Hoofring L, Jacobsen PB, Knight SJ, Learson K, Levy MH, Loscalzo MJ, Manne S, McAllister-Black R, Riba MB, Roper K, Valentine AD, Wagner LI, Zevon MA (2010) Distress management. J Natl Compr Cancer Netw 8:448–485.

Hunt SM, McKenna SP, McEwen J, Backett EM, Williams J, Papp E. A quantitative approach to perceived health status: a validation study. J Epidemiol Community Health. 1980; 34(4):281-6.

Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992; 30(6):473-83.

Further reading

Cappelleri JC, Zou KH, Bushmakin AG, Alvir JMJ, Alemayehu D, Symonds T. Patient-Reported Outcomes: Measurement, Implementation and Interpretation. Boca Raton, FL: CRC Press. 2014.
Food and Drug Administration (FDA). 2009. Guidance for industry on patient-reported outcome measures: Use in medical product development to support labelling claims. http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf

King MT. A point of minimal important difference (MID): a critique of terminology and methods. Expert Review of Pharmacoeconomics & Outcomes Research, 2011; 11(2): 171-184.
Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). http://appliedresearch.cancer.gov/pro-ctcae/