



Fiona Mellor and Karen Knapp

## Chapter Points

- Outcome measures need to be selected at the research planning stage. Selecting an outcome measure after data have been collected increases the risk of bias and type I/type II errors.
- Primary and secondary outcome measures are used to report research findings. A primary outcome measure answers the most important question and, in quantitative pilot studies, may inform the power calculation that gives rise to a sample size.
- Outcome measures are used in all methodologies (quantitative, qualitative, and mixed methods). Examples of outcome measures include imaging tests (e.g. radiographs to assess rheumatoid arthritis), physiological processes (e.g. blood pressure), performance based (e.g. sit to stand tests), and questionnaires which may be administered by a researcher or self-completed by a participant.
- Some outcome measures have data from healthy populations allowing for comparison; others allow for baseline assessments.
- Outcome measures are assessed with a variety of instruments. It is important to select an established outcome measure instrument to facilitate consistency in reporting between/across studies, and to aid in the comparisons of findings and systematic reviews/meta-analyses.

---

F. Mellor (✉)  
AECC University College, Bournemouth, UK  
e-mail: [fmellor@aecc.ac.uk](mailto:fmellor@aecc.ac.uk)

K. Knapp  
South Cloisters, University of Exeter, Exeter, UK  
e-mail: [k.m.knapp@exeter.ac.uk](mailto:k.m.knapp@exeter.ac.uk)

## 11.1 Introduction

Within research the acronym PICO (population, intervention, comparator, and outcome) lends weight to the importance of choosing an outcome measure at the initial stage of designing research. The acronym refers mainly to quantitative clinical research but in all research selecting the correct outcome is an important consideration in a design stage of research. Selecting outcome measures at a design stage is crucial in ensuring the correct order, quality and quantity of data collected. An outcome measure may also be known as a ‘construct’ or ‘domain’ [1] and they provide a common language for reporting, allowing comparisons to be made with previous research. Outcome measures are used for many reasons including the following.

- Discriminating between patients with differing disease severity at any one point in time
- Predicting patient outcome
- Measuring patient experience
- Evaluating change following an intervention
- Comparing results to normative populations
- Appraising patient safety
- Determining effectiveness of care and clinical outcomes

Different methodologies (quantitative, qualitative, and mixed methods) require different outcome measures and knowing the focus of the research is the first step in selecting an appropriate measure. Noting that ‘not everything that can be counted counts, and not everything that counts can be counted’ [2] can help target the selection of an outcome measure. For example, investigating ‘personal stress’<sup>1</sup> can be undertaken with a quantitative or qualitative approach but because there is no direct way of measuring stress a quasi-measurement/outcome measure must be selected. Within a quantitative paradigm a suitable proxy outcome measure may be blood pressure or cortisol levels, which are widely associated to rise during periods of stress [3]. A qualitative paradigm meanwhile may utilise patient-reported outcome measures (PROMs) such as interviews, focus groups, symptom status, physical function, mental health, and well-being [4]. It is beyond the scope of this chapter to discuss levels of data (nominal, ordinal, interval, and ratio, also known as numerical or categorical) but it is important to understand the classification, or level, of data to select an appropriate outcome measure. In the example of stress, the research methodology, question, aims, objectives, and research design will determine whether blood pressure and cortisol or social function and well-being, are the most appropriate outcome measures.

---

<sup>1</sup> Defined as a state of mental or emotional strain or tension resulting from adverse or demanding circumstances.

The judicious selection of outcome measures at a research design stage reduces the risk of searching for an outcome, any outcome, in the data (a type 1 error) and also provides the basis for an estimation of the sample size necessary for an adequately powered future study [5, 6]. Thus primary and secondary outcome measures should align directly with a study's aims and objectives; a primary outcome measure is often incorporated in the hypothesis.

Researchers are often tempted to develop their own outcome measures pertinent to their research question; however this approach limits the usefulness of comparing the findings to other similar studies. Outcome measures are commonly selected for inclusion in systematic reviews and meta-analyses, and existing guidelines (COSMIN, consensus-based standards for the selection of health measurement instruments) [7] exist specifically for the selection of patient-reported outcome measures (PROMs) in systematic reviews (see PROMs in Sect. 11.3 below). Researchers are therefore cautioned against developing their own outcome measure, particularly as creating a new outcome measure requires extensive evaluation to determine validity, reliability, consistency, and ability to measure the change [8] in both a statistically and clinically meaningful way. Developing new outcome measures is a branch of research in itself and beyond the scope of this chapter.

---

## 11.2 Classifications of Outcome Measures

Outcome measures may be concerned with just one aspect (uni-dimensional) such as pain, or they may be multi-dimensional and look at many facets, such as physical, emotional, and social well-being. They may relate to one condition only (disease-specific) or be more generic and relate to overall health and well-being.

Various agencies have developed taxonomies for outcome measures; many of which overlap and complement each other. Hundreds of outcome measures exist for research, patient care, and service evaluation, and these can help focus a search for a suitable outcome measure. Outcome measures for research may focus on a range of scenarios which are not exclusive to research but are pertinent when considering the primary and secondary aims and objectives. The COMET (Core Outcome Measures in Effectiveness Trials) initiative provides one such taxonomy for outcome measures which states are 'intended for the classification of what, rather than how, outcomes are measured' [9].

Wilson and Cleary [10] categorised outcome measures into five types.

1. Biological/physiological variables (e.g., ranges of motion, radiographic changes, pulse rate)
2. Symptoms status (e.g., pain)
3. Functional status (e.g., return to work)
4. General health perceptions (e.g., various aspects of global health)
5. Quality of life (e.g., general well-being, patient satisfaction)

More recently, NHS England and the Better Care Taskforce [11] developed classifications for outcome measures for service evaluation and monitoring which are also pertinent to research. These include:

1. Developing a baseline from which to measure changes.
2. Mapping and demonstrating the needs of the population being studied.
3. Monitoring clinical practice.
4. Health and disease area.
5. Target population.
6. Methodology/methods.
7. Stakeholders.
8. Study type (e.g., longitudinal/cross-sectional design).
9. Cost effectiveness and financial analysis.
10. Modelling impact.
11. Measuring quality/experience and cost.

Similarly the University of Oxford categorised outcome measures into seven domains with examples and explanations of each category [12]. These focus on PROMs and include the following.

1. Disease-specific (e.g., asthma)
2. Population specific (e.g., child health)
3. Dimension specific (e.g., Beck Depression Inventory)
4. Generic (e.g., SF-36)
5. Individualised (e.g., patient generated index)
6. Summary items (e.g., general household survey questions about long standing illness)
7. Utility measures (e.g., EQ-5D)

In 2011 the British Dietetics Association (BDA) developed a model for dietetic outcomes after recognising that many current therapy outcome measures were not generally applicable, amenable, or transferable to the work of dietitians and where the emphasis needed to be more focused on nutrition [13]. There were six domains in their classification.

1. Symptom changes
2. Physical (e.g., anthropometry/body)
3. Biochemical
4. Psychological
5. Behaviour changes
6. Patient focused

The 2011 model was superseded by a model for dietetic practice which informs education and practice and demonstrates how outcome measures can also be profession specific [14]. There are however no profession specific outcome measures for radiography due to the vast and diverse nature of this profession.

There is clearly overlap within these categories. For instance, most outcome measures can be used as a baseline (which is a domain within the NHS Better Care Taskforce taxonomy). The general health perceptions categorisation from Wilson and Cleary [10] matches with the generic criteria from the Oxford PROM group [12] and the psychological classification [13] from the British Dietetics Association (health and disease) are defined categories in both the NHS Better Care Taskforce and Oxford PROM group classifications. Some disease sites have very well-developed outcome measures, for instance within rheumatology the OMERACT initiative (Outcome Measures in Rheumatology) developed a consensus of a core set of outcome measures for rheumatology drug trials. This has now developed into an international community of health professionals with an annual conference and large database of outcome measures [15].

Developing a taxonomy for diagnostic and therapy radiography would be difficult given the breadth and depth of the profession and research across many arenas. As noted by the OMERACT group [16] the seemingly simple questions of ‘what’ and ‘how’ to measure belie a complex structure which is difficult to untangle. Thus a radiography researcher needs a thorough knowledge of the literature and research within their area of study to select appropriate outcome measures. The primary outcome measure should align closely with the stated aim of the study, and secondary outcome measures should align to the objectives.

---

### 11.3 Patient Reported Outcome Measures (PROMs)

Patient-reported outcome measures (PROMs) are essentially any outcome directly reported by a patient. They are generally subjective in nature and range from simple visual analogue scales (VAS), used to report pain intensity on a scale of 1–10, to more in-depth questionnaires completed over a lifetime of a longitudinal study. They are typically used to collect patients’ perceptions and views about their health, health status, quality of life, and care.

Health research used to be dominated by outcome measures selected by researchers without acknowledgement of their bias (a surgeon may be more likely to select a surgical outcome measure such as blood loss rather than an outcome measure of importance to a patient such as quality of life). However the inclusion and acceptance of patient and public involvement (PPI) in research, along with the phrase ‘nothing about us without us’ pushed PROMs to the fore. Patient reported outcome measures ensure that what matters to a population under investigation is included in the results and can be used to influence future decisions and policies and there is now a trend in healthcare research to use PROMs to focus more on patient-centred research [17, 18]. They can also act as a quality improvement strategy for patient care through feedback; it has however been noted that in some cases, such as palliative care and psychotherapy, clinicians viewed individualised PROMs as more useful to build rapport rather than substantially change communication practices [19].

Patient reported outcome measures became popular in the NHS in 2009 following Lord Darzi’s 2008 report *High quality care for all* [20]. Since then the NHS has collected PROMs in four surgical procedures: hip replacement, knee replacement,

varicose vein surgery, and hernia surgery, and they publish an annual web based report [21] and an annual national conference on the development and use of PROMs [22]. The use of PROMs in research ensures that important knowledge about the impact of an intervention is not lost because the selected measure was unable to capture it or, even worse, distorted the true results [23].

It is essential to consider the use of PROMs in research to demonstrate impact. Such measures (including quality of life and symptoms), if collected, analysed, and reported appropriately, can be used to inform shared decision-making, clinical guidelines, and health policy. However one problem with interpreting PROMs is distinguishing how much movement on a scale equates to a clinically meaningful change. While scores on scales can be subject to rigorous statistical analysis, a statistically significant difference may not equate to a clinically meaningful change. A statistical test is used to determine whether an effect is likely to be due to chance or not and if a study is sufficiently powered then a small change can be statistically significant but this does not mean it is clinically significant. For many physiological measurements (such as temperature or blood pressure), experience and clinical judgement inform whether the results are clinically meaningful but with more subjective measurements, such as pain or stress, it becomes harder to define [24].

Patient reported outcome measures were recently sub-classified into patient satisfaction measures and patient-reported experience measures (PREMs) [25] and are discussed below in Sect. 11.3.2.

### 11.3.1 Patient Reported Experience Measures (PREMs)

Patient reported experience measures (PREMs) differ from PROMs in that the former capture participants' perception of their experience (with their healthcare or service) whilst PROMs focus on participants' perception of their health. Examples of PREMs include quality of communication, time spent waiting, and whether they would recommend the service to family and friends (the NHS Friends and Family test) [28].

In England, the Department of Health [20] defines the three domains of care as: patient safety, clinical effectiveness, and patient experience; commissioners and service providers are increasingly using PREMs in their assessments. A study in 2014 noted that there was a weak positive correlation between PROMs and PREMs in elective surgery in the UK; the authors [29] stated that PREMs may not be used as a proxy for a good outcome. Thus caution is advised in the use of PREMs as a primary outcome measure if the objectives of a study are not directly related to patient experience.

### 11.3.2 Patient Satisfaction

Patient satisfaction may be both a PROM or a PREM and are frequently used to enable researchers to focus on the impact of their research.

Measures of satisfaction differ from outcome measures such as quality of life (QoL); they address the process of treatment rather than its outcome [26]. However, patient satisfaction scales are sometimes referred to as ‘happy scales’ because they can mask negative experiences. Patients tend to score their care highly and there is little discrimination between items which can lead to a noted ceiling effect<sup>2</sup> [25]. A study of 21 EU countries [27] indicated that satisfaction is also linked to ‘broader societal factors’ such as the wealth and prosperity of a country.

---

## 11.4 Considerations for Choosing an Outcome Measure

Pertinent outcome measures are identified at the same time as the literature review, which is undertaken to justify a research question. However consulted research may utilise a number of different outcome measures and it may be difficult to choose the most useful way of measuring the outcome. The challenge is to ensure selected outcome measures truly reflect the change intended to be assessed. For example a scale measuring pain may be broad enough to include both acute and chronic pain, or focused enough to measure just one aspect of pain. If you are interested purely in comparing the efficacy of one intervention to manage immediate post-operative pain then a scale that measures both chronic and acute pain would likely to be misleading, as would a scale measuring chronic pain only.

With respect to PROMs both barriers and facilitators to their implementation have both been examined, and specifically in palliative care their routine use has been slow and difficult. An educational component (e.g. understanding how to complete and score PROMs) alongside an understanding of the emotional and cognitive processes of a patient, were deemed as crucial in this area of health [30].

A 2016 Delphi study [1] of 120 participants developed guidelines for selecting a ‘core outcome set’ (COS) of outcome measurement instruments (OMIs) to be reported in all clinical trials of a specific disease or trial population. This was in co-operation with the COMET initiative [9] which has an online searchable database of outcome measures and a wealth of advice on how to select the most appropriate outcome measure. Specific considerations for selecting outcome measures are presented below.

---

## 11.5 Licencing and Costs

Some well-known outcome measures are free to use; many others incur a charge. Outcome measures are copyrighted and may have a licence with associated terms and conditions which a researcher needs to be aware of. Whilst it may be possible to

---

<sup>2</sup>A ceiling effect is when the top scale on the measurement instrument is consistently reached, thus reducing the ability of the scale to accurately capture data beyond the top of the scale.

develop a similar outcome measure with new wording it is important to bear in mind that reliability and validity are based on exact wording. Consequently changing the structure would mean the new outcome measure has no evidence to support its use and it would be difficult to prove that it consistently measures a result as expected. It is recommended that a developed and established outcome measure is used as intended and that licensing conditions are checked and adhered to before research begins.

---

## 11.6 Population and Stakeholders

In research a population is defined as a collection of individuals or objects with similar characteristics and they may be categorised by disease type, profession, or demographics. When selecting an outcome measure it is important to establish whether it has been used in a population similar to the one under consideration. Questions to ask include: Has the outcome measure been used in the same condition and disease severity as the sample being studied? Is it responsive to the differences you hope to detect? If level of pain is the primary outcome measure is acute or chronic pain being measured? Is a patient or clinician reporting the results?

Stakeholders by contrast are consumers of research outcomes and include patients, providers, payers, regulators, industry, academic, society, and policy-makers [1]. The increasing acknowledgment of stakeholders in research is one driver for using PROMs. Orthopaedic research interest in PROMs dates back to the 1980s [31] but useage in this and other clinical areas has been slow. It is still usual to see clinical outcome measures that focus on the technical success of an operation such as mortality, morbidity, and complications, as opposed to patient satisfaction and quality of life.

A population and stakeholders may have different priorities in terms of what constitutes meaningful information from research and this needs to be accounted for in the selection of primary and secondary outcome measures. It is worth noting that PROMs need to be meaningful to many groups including patients, clinicians and researchers [32].

---

## 11.7 Administering and Scoring Results

It is essential to ensure a selected outcome measure is acceptable to both researcher and participant. It should be quick and simple to use, reliable, valid, specific to the question being investigated, and cost-effective. Longer outcome measures may collect full and comprehensive data but will demand a greater input from a respondent and a more thorough analysis beyond the scope of a research project. Conversely basic or simple outcome measures may not provide enough information to quantify results and report differences within or between groups.



The length of an outcome measure is important when considering a population under study. For instance, paediatric or brain injury research requires short and easy understandable outcome measures if they are to be completed by a patient (PROMs). Equally clinical outcome measures need to be short and easy to complete if a large sample size in a short period of time is under investigation. Conversely, and with respect to the study design, a longitudinal study with a smaller sample size may require an established but complicated outcome measure, which could be more easily justified than in a cross-sectional case study design.

Whilst some outcome measures may be relatively easy and simple to administer, scoring and reporting may be more complicated and may require the use of complex statistical packages. An established outcome measure should include clear standardised instructions on how to implement and score the results and whether specialist training may be required, in which case this will need to be costed within the research.

---

## 11.8 Responsiveness

The responsiveness of an outcome measure relates to how well it can detect a change over time (longitudinal) or differences between groups. Responsiveness applies to the efficacy and effectiveness of an intervention [33]. For a measure to be responsive it also needs to be reliable and valid. The characteristics of an outcome measure affect its responsiveness, for example, ordinal data (data that place participants into categories such as ‘constant’, ‘frequent’, ‘occasional’, ‘rare’, or ‘never’) are likely to be less responsive than interval or ratio data (data on an established scale such as temperature) if large changes in status are required to change categories.

This consideration needs to be balanced against the amount of data collected and the time needed to analyse the data. Additionally outcome measures with ceiling or floor effects (that do not account for improvement or decline, or where the baseline includes a very high or very low score) may not be responsive. An example would be using an activity outcome measure which assesses level of assistance needed for certain tasks elderly post-surgical patients versus young amputees being trained for high level sports, would give a very different outcome in both groups [34].

---

## 11.9 Reliability and Validity

Reliability refers to how effectively an outcome measure can be repeated on different occasions with the same conditions and provide the same result. It is essentially the consistency of a measurement. Conversely validity refers to how well an outcome measure can assess a feature being measured. A simple explanation of reliability and validity would be an alarm clock set for 7.00 am every morning but which consistently rings at 6.30 am. IT would be a very reliable alarm clock, but it would not be valid [35].

There are different types of reliability and validity: both are divided into three sub-domains.

Reliability sub-domains include the following.

- Internal consistency is the extent that all items on a scale measure the same thing. For instance, does a pain scale only measure pain or does it also include questions which measure disability?
- Stability is the consistency of results in repeated testing in the same and different populations.
- Equivalence is the level of agreement between the interpretations of the scores; sometimes known as intra- and inter-rater agreement.

Validity sub-domains include the following.

- Content validity is how well a measure captures all the features of the domain. For instance, if both acute and chronic pain are being measured does the intended outcome measure cover both aspects or only acute pain?
- Construct validity is how well an intended outcome is measured. For instance, does a pain measurement only allow a score of 1–10 on a visual analogue scale (VAS) or does it also allow for responses measuring the use of painkillers?
- Criterion validity is how well a measure relates to other measures which examine the same outcome. For instance, does a VAS for disability relate to other disability outcome measures? [36].

Both reliability and validity of a chosen outcome measure should be noted when choosing the most appropriate way to record and interpret the results of a research. Although these are not always recorded it is imperative that they are investigated if a new outcome measure is being developed.

---

## 11.10 Subjective Versus Objective Outcome Measures

Outcome measures can be subjective or objective. Both have challenges in practice and there is no ‘one size fits all’ approach. An objective measure is more likely to consider medical data and be collected by professional equipment such as pulse oximeters, cardiac monitors, or biochemistry data. Such measures are precise and reduce bias thus they are widely used in research. They tend to collect short-term data which can change quickly.

Subjective measurements are defined as those which are open to interpretation such as questionnaires and visual analogue scales (VAS). They are generally quick and easy to administer and can be completed by a researcher or patient but they are more prone to bias and errors in both completing and reporting results. Pain as an outcome measure is subjective as demonstrated in a study of nursing students who were asked to identify pain on a VAS [37]. They reported a wide range of perceptions demonstrating the highly subjective nature and different terminology associated with pain. Radiographic scoring methods, often used in an assessment of onset

and progression of joint disease/degeneration, are also subjective. Numerous methods for this exist [38] and the results are often statistically reported because their scales and data distribution meet parametric properties and standards. To reduce bias in such cases it is preferable to have more than one independent assessor to enable intra- and inter-rater agreement to be presented [39].

---

### 11.11 Confounders

It is important to consider extraneous uncontrolled independent variables, or confounders, when designing a study and selecting an outcome measure. A confounder is a distortion of an association of an intervention and outcome. Unlike other kinds of research bias, such as selection or researcher bias, a confounder can be adjusted in an analysis providing the researcher knows the confounding variables in advance. Such variables often include patient demographics, such as height, weight, age, gender, race, but might also include a multitude of other variables such as stage of disease, co-morbidity, signs and symptoms, duration of disease, and types of imaging equipment used [40]. For example, bone density reduces with age in normal and osteoporotic subjects. If an outcome measure is the odds ratio and there is no adjustment for age then this will lead to an over-estimation of the discrimination of fracture cases from controls. Therefore age adjustment is required in the analysis to correct the problem [39].

---

### 11.12 Searching for an Outcome Measure

It is useful to search relevant literature and systematic reviews to identify an appropriate outcome measure. There are also a number of online searchable databases for outcome measures although these are subject to change with little or no notice. As with literature it is worth searching a number of databases to identify more popular outcome measures. When searching online databases take careful note of the owner of the URL and also the number of studies registered that have used, or are using, the outcome measure of interest.

The databases presented here are current at the time of writing.

#### 11.12.1 The COMET Initiative

The COMET initiative focuses on effectiveness trials and brings together a core outcome set (COS) of minimum outcome measures which should be reported in clinical trials of a specific condition. It also explains problems and key issues in using outcome measures in research. This database (<http://www.comet-initiative.org/>) is endorsed by the European Commission, The Medical Research Council (MRC), The National Institute for Health Research (NIHR), and the Seventh Framework Programme.

### 11.12.2 Rehabilitation Measures Database

This is a privately held and free to search database of outcome measures associated with rehabilitation which are predominantly questionnaire based. This database (<https://www.sralab.org/rehabilitation-measures>) gives an overview of the listed outcome measures including the time expected to complete and costs of the licence. It is organised by assessment type (including PROMS and performance measures), area of assessment including activities of daily living (ADLs), patient satisfaction, and populations including allied health care professions, joint care, and fractures.

### 11.12.3 Patient Reported Outcome Measurement Group

This PROM group is based at the University of Oxford. The website (<http://phi.uhce.ox.ac.uk/home.php>) includes guidance on the selection of appropriate outcome measures for use in clinical trials, practice, and population surveys. It is no longer supported with updates but it is still a useful resource.

### 11.12.4 Proqolid

This database (<https://eprovide.mapi-trust.org/about/about-proqolid>) [41] provides information on over 2000 clinical outcome assessment tools. There are two levels of access; free and subscription. The free version provides basic information on outcome measures including the therapeutic area, indication, and bibliographic references for the description of the outcome measure.

### 11.12.5 OMERACT

Outcome Measures in Rheumatology (OMERACT) is an online database of outcome measures pertinent to rheumatoid diseases (<https://omeract.org/>) as well as clinical and radiographic outcome measures [42]. There is a heavy emphasis on patient involvement in endorsing outcome measures and advice on how to select an appropriate outcome measure.

---

## 11.13 Reporting Outcome Measures

Both quantitative and qualitative research follow a structured approach for writing and reporting results and the choice and use of outcome measures need to be clearly stated in research papers. A good example is presented by Hardy et al. [43].

The methods section acts as a recipe and provides details to allow full replication. Within this section it is important to clearly state the outcomes measured and the tools used for measurement. The primary outcome measure needs to align with

the overall aim of a study and secondary outcome measures align with its objectives. It is important to provide enough detail for a reader to know exactly what was measured and how.

The results section displays the outcomes of research and many outcome measures may have standard criteria for this. Following the stated criteria enables comparison with other research that used the same outcome measures.

The discussion section is where the results are compared to other research and the comparison is easier if other research has used the same outcome measures. It is also the section where the choice of outcome measures can be justified. This includes providing substantiated evidence of the outcome measures' reliability and validity, and relating these to the primary and secondary outcome measures with evidence of how they collectively support a study's aims. There are many instances where more than one outcome measure would have captured the relevant data but including too many outcome measures can lead to an unfocused research question and present problems with interpretation if the effect differs across the outcomes [44].

The conclusions of the study state the composite endpoints which are the amalgamation of the selected outcome measures that have been correlated to support or refute the hypothesis. The components of selected outcome measures should be complementary, with secondary outcome measures lending supporting evidence to the primary outcome measure.

---

## 11.14 Outcome Measures in Radiography Research

There are no recommendations for specific outcome measures for radiography research because the profession spans a range of conditions and diseases, and research within radiography is varied and far ranging.

However radiographic scoring methods are a well-established outcome measure in many diseases including rheumatology [38] and PROMs are frequently used in cancer research to establish the impact of cancer treatment [45]. A systematic review in 2006 found there were no PROMs developed specifically for radiology [46].

A focus of research within imaging and oncology is new techniques and technology development which are regularly introduced and require evidence to test their ability to detect pathologies or to predict clinical outcomes, cost effectiveness, and patient-centred outcomes. Such research often uses tests of reliability and accuracy as outcome measures, including sensitivity and specificity.

If there is already a technique for the diagnosis of a pathology of interest it is important to directly compare the techniques. The best method is to undertake this within the same patients so they would have two imaging or diagnostic techniques rather than one, however this is not always practicable or possible. There is also the potential that a combination of a new and old technique can improve the diagnostic accuracy above and beyond either technique when used individually, in which case this needs to be included in the study design and analysis.

Secondary outcome measures of new technology research often include the radiation dose associated with each of the techniques, the cost of the technique in terms of duration, cost of the equipment and consumables, how invasive the technique is, and consumers' perception with the treatment (PROMs). While some lower dose and cheaper techniques may be introduced it is important that they match or exceed the diagnostic accuracy of an existing gold standard [39].

In radiotherapy outcome measures are used extensively in quality assurance such as cancer wait times with statistics widely available from the NHS website [47]. This outcome measure does not alone capture the depth and breadth of radiotherapy and is rarely used in cancer research studies. Many radiographers may be involved in cancer research and may work within a research network ensuring the quality assurance of radiotherapy trials. Within the UK the Radiotherapy Trials Quality Assurance group (RTTQA) ensures all radiotherapy trials are conducted to the same standard. Whilst the RTTQA group is not concerned with outcome measures per se, it does hold the full trial protocols of radiotherapy studies and these protocols include the selection and justification of the outcome measures selected.

An example of such a trial called Fast Forward, which is a trial of radiotherapy in breast cancer, lists the primary outcome measure as ipsilateral local tumour control and secondary outcome measures as early and late adverse effects in normal tissues, quality of life, contralateral primary tumours, regional and distant metastases, and survival and publications regarding skin toxicity (a secondary outcome measure) have already begun to influence practice [48]. Quality of life is the only PROM in this study because its primary objective is to identify whether a five fraction schedule of curative radiotherapy is at least as effective and safe as the 15 fraction regime, but other breast cancer studies have centred on PROMs such as the START trial that examined patient reported breast, arm, shoulder symptoms, and body image [49].

Whilst there are no recommended outcome measures for radiography research, the Society and College of Radiographers (SCoR) have set research priorities which fall under four domains: accuracy and safety; effectiveness of technical approaches; the patient experience; and service delivery and organisation [50]. Aligning radiography research aims and objectives within these priorities would aid in selecting the most suitable primary and secondary outcome measures.

---

## 11.15 Conclusion

The final take home message is that outcome measures are a central aspect of research and need to be considered from the initial design stage. It is important to take time to select the most appropriate outcome measures, taking into account the considerations for choosing an outcome measure as covered in this chapter. Finally researchers need to justify their choice of outcome measure(s) and ensure they align with the aims and objectives of their research.

## References

1. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” a practical guideline. *BMC Trials*. 2016;17(1):449.
2. Cameron WB. *Informal sociology a casual introduction to sociological thinking*. New York: Random House; 1963.
3. Rainforth MV, Schneider RH, Nidich SI, Gaylord-King C, Salerno JW, Anderson JW. Stress reduction programs in patients with elevated blood pressure: a systematic review and meta-analysis. *Curr Hypertens Rep*. 2007;9(6):520–8.
4. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. *BMJ (Clinical research ed)*. 2015;350.
5. Andrade C. The primary outcome measure and its importance in clinical trials. *J Clin Psychiatry*. 2015;76(10):1320–3.
6. Ferreira JC, Patiño CM. Types of outcomes in clinical research. *J Bras Pneumol*. 2017;43(1):5.
7. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, HCW DV, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147–57.
8. Mohtadi NG. Outcome measure development. *Instr Course Lect*. 2016;65:577–82.
9. Dodd SR, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol*. 2018;96:84–92.
10. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *J Am Med Assoc*. 1995;273(1):59–65.
11. NHS Better Care Fund Taskforce. ‘How to guide’; the better care taskforce technical toolkit. Section 3. Outcomes and impact measurement. London; 2014.
12. University of Oxford. Patient reported outcome measures: instrument types. Oxford; 2010 [cited 2019 May 13]. [http://phi.uhce.ox.ac.uk/inst\\_types.php](http://phi.uhce.ox.ac.uk/inst_types.php).
13. The Association of UK Dietitians. Model for dietetic outcomes. Birmingham; 2011 [cited 2019 May 15]. [https://www.bda.uk.com/publications/archive/bda\\_outcome\\_model\\_2011\\_archive](https://www.bda.uk.com/publications/archive/bda_outcome_model_2011_archive).
14. The Association of UK Dietitians. Model and process for nutrition and dietetic practice. 2017 [cited 2019 May 14]. <https://www.bda.uk.com/professional/practice/process>.
15. OMERACT (Outcome Measures in Rheumatology) [cited 2019 May 20]. <https://omeract.org/>.
16. Boers M. How outcome measures in rheumatoid arthritis clinical trials works to develop outcome measures in rheumatology clinical trials. Choosing and developing core outcome measurement sets for clinical trials: outcome measures in rheumatoid arthritis clinical trials filter 2.0. *Rheumatology*. 2017;56:kex060.116.
17. Speight J, Barendse SM. FDA guidance on patient reported outcomes. *BMJ*. 2010;c2921:340.
18. Staniszevska S, Haywood KL, Brett J, Tutton L. Patient and public involvement in patient-reported outcome measures: evolution not revolution. *Patient*. 2012;5(2):79–87.
19. Greenhalgh J, Dalkin S, Gooding K, Gibbons E, Wright J, Meads D, Health Services and Delivery Research, et al. *Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care*. Southampton: NIHR Journals Library; 2017.
20. Darzi A. *High quality care for all: NHS next stage review final report*. London: The Stationery Office; 2008.
21. NHS Digital. Patient reported outcome measures (PROMs). 2019 [cited 2019 May 15]. <https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms>.
22. Gibbons E, Calvert M, Bostock J, Skryban M. Proceedings of patient reported outcome measures (PROMs) conference Birmingham 2018: Birmingham, UK. 20 June, 2018. *J Patient Rep Outcomes*. 2018;2(Suppl 2):58.



23. Coster WJ. Making the best match: selecting outcome measures for clinical trials and outcome studies. *Am J Occup Ther.* 2013;67(2):162–70.
24. Bolton JE. Sensitivity and specificity of outcome measures in patients with neck pain: detecting clinically significant improvement. *Spine.* 2004;29(21):2410–7; discussion 8.
25. Hodson M, Andrew S, Michael Roberts C. Towards an understanding of PREMS and PROMS in COPD. *Breathe.* 2013;9:358–64.
26. McKenna SP. Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science. *BMC Med.* 2011;9(1):86.
27. Bleich SN, Ozaltin E, Murray CJL. How does satisfaction with the health-care system relate to patient experience. *Bull World Health Organ.* 2009;87:271–8.
28. National Health Service. Friends and family test. 2018 [cited 2019 May 20]. <https://www.nhs.uk/using-the-nhs/about-the-nhs/friends-and-family-test-fft/>.
29. Black N, Varaganum M, Hutchings A. Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery. *BMJ Qual Saf.* 2014;23:534–42.
30. Antunes B, Harding R, Higginson IJ. Implementing patient-reported outcome measures in palliative care clinical practice: a systematic review of facilitators and barriers. *Palliat Med.* 2014;28(2):158–75.
31. Brinker M, O'Connor D. Stakeholders in outcome measures: review from a clinical perspective. *Clin Orthop Relat Res.* 2013;471(11):3426–36.
32. Duncan E, Murray J. The barriers and facilitators to routine outcome measurement by allied health professionals in practice: a systematic review. *BMC Health Serv Res.* 2012;12(1):96.
33. Tarrant C, Angell E, Baker R, Boulton M, Freeman G, Wilkie P, et al. Responsiveness of primary care services: development of a patient-report measure – qualitative study and initial quantitative pilot testing. *Health Serv Deliv Res.* 2014;2(46):1.
34. Roach KE. Measurement of health outcomes: reliability, validity and responsiveness. *J Prosthet Orthot.* 2006;18(6):P8–P12.
35. Heale R, Twycross A. Validity and reliability in quantitative studies. *Evid Based Nurs.* 2015;18(3):66.
36. Boonstra AM, Schiphorst Preuper HR, Reneman MF, Posthumus JB, Stewart RE. Reliability and validity of the visual analogue scale for disability in patients with chronic musculoskeletal pain. *Int J Rehabil Res.* 2008;31(2):165–9.
37. Bergh I, Sjostrom B. Quantification of the pain terms hurt, ache and pain among nursing students. *Scand J Caring Sci.* 2007;21(2):163–8.
38. Boini S, Guillemin F. Radiographic scoring methods as outcome measures in rheumatoid arthritis: properties and advantages. *Ann Rheum Dis.* 2001;60(9):817–27.
39. Mellor F, Knapp K. Research outcome measures. In: Ramlal A, editor. *Medical imaging and radiotherapy research: skills and strategies.* London: Churchill Livingstone; 2010. p. 111–21.
40. Crewson PE, Applegate KE. Data collection in radiology research. *Am J Roentgenol.* 2001;177(4):755–61.
41. Pinotti R. PROQOLID. *J Med Libr Assoc.* 2016;104(1):91–2.
42. Kelly A, Tong A, Tymms K, March L, Craig JC, De Vera M, et al. Outcome measures in rheumatology - core domain set for trials of interventions for medication adherence in rheumatology: 5 phase study protocol. *Trials.* 2018;19:204.
43. Hardy M, Hutton J, Snaith B. Is a radiographer led immediate reporting service for emergency department referrals a cost effective initiative? *Radiography.* 2013;19(1):23–7.
44. Vetter TR, Mascha EJ. Defining the primary outcomes and justifying secondary outcomes of a study: usually, the fewer, the better. *Anesth Analg.* 2017;125(2):678–81.
45. Faithfull S, Lemanska A, Chen T. Patient-reported outcome measures in radiotherapy: clinical advances and research opportunities in measurement for survivorship. *Clin Oncol.* 2015;27(11):679–85.
46. Mathers SA, Chesson RA, Proctor JM, GA MK, Robertson E. The use of patient-centered outcome measures in radiology: a systematic review. *Acad Radiol.* 2006;13(11):1394–404.



47. National Health Service. Cancer waiting times. [cited 2019 May 15]. <https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/>.
48. Brunt AM, Wheatley D, Yarnold J, Somaiah N, Kelly S, Harnett A, et al. Acute skin toxicity associated with a 1-week schedule of whole breast radiotherapy compared with a standard 3-week regimen delivered in the UK FAST-forward trial. *Radiother Oncol.* 2016;120(1):114–8.
49. Hopwood P, Haviland JS, Sumo G, Mills J, Bliss JM, Yarnold JR. Comparison of patient-reported breast, arm, and shoulder symptoms and body image after radiotherapy for early breast cancer: 5-year follow-up in the randomised standardisation of breast radiotherapy (START) trials. *Lancet Oncol.* 2010;11(3):231–40.
50. Society and College of Radiographers. Society and College of Radiographers Research Priorities. [cited 2019 May 20]. [https://www.sor.org/system/files/article/201312/scor\\_research\\_priorities.pdf](https://www.sor.org/system/files/article/201312/scor_research_priorities.pdf).