Patient Reported Outcomes Measures in radiotherapy: clinical advances and research opportunities in measurement for survivorship

Prof. Sara Faithfull
School of Health and Social Care, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7TE, UK

Dr Agnieszka Lemanska
School of Health and Social Care, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7TE, UK

Dr Tao Chen
Department of Chemical and Process Engineering, University of Surrey, Guildford, GU2 7XH, UK

Abstract
Patient reported outcome measures (PROMs) are a useful way of recording patient perceptions of the impact of their cancer and the consequences of treatment. Understanding the impact of radiotherapy longer term requires tools that are sensitive to change but also meaningful for patients. PROMs are useful in defining symptom severity but also the burden of illness for cancer patients. Patient reported outcomes are increasingly being seen as a way to improve practice by enhancing communication, improving symptom management as well as identifying patient care needs. This paper provides an overview of the use of PROMs in radiotherapy and considerations for tool choice, analysis and the logistics of routine data collection. Consistent assessment is essential to detect patient problems as a result of radiotherapy but also to address emerging symptoms promptly.

Introduction
The number of cancer survivors is increasing as earlier detection and better therapies have brought about significant advances in survival rates. In Europe 46% of all those diagnosed with cancer will be living 10 or more years beyond initial treatment (1). By 2030 it is projected that there will be more than 4 million cancer survivors within the UK population (2) and 13.7 million in the USA (3). Monitoring the health of survivors is crucial in comparing long-term outcomes from cancer therapies (4) (5). Recording the impact of different treatments on individuals quality of life and health burden is an important part of patient reported outcomes (6),(7).

Systematic monitoring of late effects from radiotherapy is important in determining normal tissue effects and measuring dose response (8). However, consistent assessment is rarely undertaken routinely in the UK. It is however essential to detect patient problems as a result of therapy but also to address emerging symptoms promptly and meet patients healthcare needs (9, 10). Consideration beyond treatment efficacy particularly with determining potential risk of long term side-effects is required for informing clinician and patient decision making, especially when differences in survival maybe small between therapies (11).

Self-rated health measures have been found to be a powerful predictor of morbidity and mortality compared to many objective measures of health (Chase et al 2012) so show
evidence of substantial benefit for clinical practice (12). Patient reported measures have been shown to have equivalence to physician reported measures in radiotherapy (13, 14). Developing prospective and long-term methods for future patient assessment requires measurement of patients’ perceptions of their symptoms and recording of health outcomes over a considerable period of time.

Patient reported outcomes (PRO) are currently a force for clinical improvement driven not only by health policy initiatives but also to improve patient centred health (15), (16). 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” (17). PROMS are increasingly being used to make comparisons of health outcomes across healthcare settings (18, 19).

This paper explores the benefits and challenges of PROMS for use in radiotherapy practice. We consider the conceptual differences between tools together with the most useful way of capturing, analysing PROMs and utilising the data. Finally we provide guidance for radiotherapy clinicians keen to use patient reported outcome measures in radiotherapy research and practice.

What are the differences between patient reported outcome measures?

There has been a proliferation of PROMS over the last few years and many are used within oncology. PROMS can range from multi-dimensional measures of patient’s global perceptions of their health to specific tools that assess severity of symptoms (16, 20). It is important to distinguish between types of PROMs as they measure conceptually different items (Figure 1).

Generic PROMs, such as Health Related Quality of Life (EQ-5D), measure the patient’s perceptions and societal values of the impact of disease and treatment. These generic tools measure health as the ability to function, often emotionally, physically and socially but can be strongly influenced by environmental factors (20). They provide population based data that is useful for comparison, provides data for health economics but have low sensitivity to change at an individual level. Disease specific PROMs ask patients about condition specific problems such as quality of life (EORTC-QOL C30) with disease specific attributes and even more specific symptom scores such as anxiety and depression and body image scale (HADs, BIS). These condition specific measures have high patient relevance and sensitivity but provide poorer population relevant data (19). Such symptom focused PROMs are best at recording severity but are prone to response shift over time for example patients adapt to health changes and therefore the impact on patient reported outcome reduces over time (21). This can be seen in Mukesh et al (22) study where PROMs reporting breast changes after radiotherapy were under reported over time compared to clinician assessments.

PROMS do not ask about patient experience of care, opinions or satisfaction with health care: these are not health outcomes (15). PROMS can provide descriptive richness and detail, they fall across a continuum of sensitivity to generalizability and the choice of measurement instrument therefore depends on the target coverage of content, sensitivity of that target to change and the potential for comparison. Combining measures to target different aspects of patient reported outcomes is an approach that enables relationships to
be explored between more global and specific impact and may be especially informative in intervention evaluation (19). To understand change PROMs must be used across time, such as before and after radiotherapy so that it is possible to gauge what improvement or deterioration has occurred and patient needs.

**Evidence of benefits of PROMs for improving clinical practice**

Increasingly patient follow-up after radiotherapy is changing with workflow alterations for clinical oncologists reducing patient contact for those after treatment and greater use of remote and telephone follow up and discharge to primary care (23). The ability to assess patients remotely and provide systematic assessment of radiotherapy late effects is reliant now on PROMs. The benefits of PROMs as part of clinical review have been observed in a range of studies. However in a meta-analysis of PROMs effect on outcomes and processes of care they were seen to increase detection of patient problems, enhance symptom control and improve supportive care measures (12). Communication was not enhanced in this meta-analysis despite indications from single studies (24). The effect size of such changes was small and the psychometric robustness of some of the PROM tools questioned by authors in their ability to evaluate the intervention required. Although routine collection of PROMS are perceived positively the need for automated systems to identify any areas that require specific attention is required to support clinician decision making (25). The logistics of collecting PROs in clinical practice is challenging and are dependent on the resources available. Useful methodological and logistical tips for successful integration of PROMS into routine clinical practice are provided (15, 26). Research is on-going in evaluating such approaches in oncology (27). PROMs are useful to facilitate open discussions, and also to identify areas of concern during treatment follow-ups.

**How are PROMs used in radiotherapy research?**

PROMs have widespread use in pharmaceutical intervention studies with detailed guidance in their inclusion to support efficacy claims and approval of new drugs. The FDA revised guidance in 2009 raised the stakes for the use of PROMs in clinical research resulting in a major shift in how PROM assessments were viewed (28). Fundamental to this was bringing the patient perspectives into developing PRO instruments emphasising the content validity and requiring that PRO instruments adequately measure what is claimed, that is they are psychometrically tested for that population and that the PRO endpoint is clear in the clinical trial. In the past such requirements for inclusion of PROs in radiotherapy research were not required however much can be learnt from standards of PRO use in identifying tool validity within radiotherapy (29) (20). Subsequently the field of PROM development is rapidly changing with tools being revised and developed constantly.

This element of change creates difficulties in comparing old with new PROMs. However this is less of a problem as item analysis and detail of analysis in reporting can address comparisons for future studies. More critically the small incidence of severe late effects and lack of specificity to radiotherapy effects mean that PROM tools rarely have precision in detecting significant differences between radiotherapy treatments. The evolving basis of treatment and emerging late effects for cancer survivors requires researchers to look ahead so that emerging treatment effects can be captured in relation the HRQOL(30). This means
there is a need for PRO measures across the spectrum (Figure 1) to be able to predict future requirements and analyse treatment effects.

Graff in 2002 (31) reviewed PRO measures used in radiotherapy between 1990 and 2001. The tools most frequently used were EORTC QLQ-C30 and the FACT G with corresponding tumour modules. We have conducted a similar analysis of studies. Table 1 gives a sample of PROM instruments used in clinical trials, which involve radiotherapy, and were either primarily conducted in the UK or participated by UK centres. Although this is not an exhaustive list, it shows that the questionnaires developed by the EORTC’s Quality of Life group are widely used, including the core, generic tool (QLQ-C30) as well as cancer specific modules. Most studies used a combination of generic and specific questionnaires, both within the EORTC tools (e.g. EORTC QLQ-C30 combined with BN20 for brain cancer) and beyond (e.g. SF-36, a generic tool combined with prostate-specific UCLA-PCI for prostate cancer), though this is not always the case. It is worth noting that when multiple tools are used together, they may contain overlapping questions, differences in scaling and these may result in differences in PRO results within the same study (32). Our results were similar to Graff in that it is important to know what you intend to measure, why you want to measure it and be sure the PRO instrument you choose measures what is intended or use tools that include both global and specific PROM concepts.

**PROMS, data analysis the challenge of incomplete data**

The general challenges that affect PROMs, just as with any other type of data, include dealing with missing data, the abundance of information generated from PROMs collected over time and the need to summarise or transform scores to reverse questions that may have different polarity i.e. negative to positive. This may require variable selection prior to further analysis and requires a clear PROM analysis plan from the beginning of the research. Most PROM tools provide guidance on analysis. In addition, depending on the study research question, stratification of patients may require attention when analysing PROM data. For example in large multicentre studies the information on hospital, deprivation or cancer stage (risk group) may need to be included. In this type of analysis patients are stratified within their groups. Groupings between patients may significantly affect the outcome of the study (and mask the main effect) and therefore it is important to take it into account to enable detection of the main effect and to comprehensively answer the experimental question.

Missing data cannot be completely prevented, due to various reasons. Often patients do not provide all the answers to the questions, and there may be specific questions that patients may not understand, prefer not to answer or may consider to be irrelevant. Alternatively, it is possible that clinicians may fail to give out or collect forms. Respective of which data are missing, missing data can be divided into two types; drop-outs and intermittent (33). Dropouts are when participants fail to deliver the whole assessment on a specified occasion (for example did not attend) and this is especially typical for longitudinal studies when some of the participants may be lost to follow-up.
One of the most commonly applied methods of dealing with missing data in PROMs is complete case analysis. It is a traditional approach that is currently not recommended in the methodological literature (34). This approach clears the data from that missing by removing incomplete cases, and this enables the statistical software to process the data. However, it can have some serious disadvantages in introducing bias. Some analytical approaches, such as those that use correlation coefficient, allow for pairwise deletion of missing variables. This preserves some data but has implications when interpreting the results because each measurement is based on a different number of samples. However, for most of the multivariate approaches e.g. regression, the whole assessment has to be deleted (listwise deletion) if it contains at least one unanswered question (35). Subsequently, there is a risk that conclusions may be biased towards patients with complete assessments, and secondly the loss of power of the analysis that is associated with the sample number may be significant. The assumption that the complete cases are random selections of samples and that they will provide adequate representation of the entire dataset is very rarely supported (36). Alternative methods are available and recommended for use in PROMs. They involve various way of imputing missing data. Traditionally, missing data in PROMs was imputed with the rounded mean or median of the variable, or in the case of dropouts by the last observation that is carried forward. However, more advanced imputation algorithms are currently being developed and used more in PROMs analysis (ref). They utilize maximum likelihood estimation and various multiple imputation algorithms. The choice of the imputation approach depends on the type of data, and has to be carefully considered taking into account mechanisms of missing data as defined by Ruben (34). Multiple imputation has been shown to be beneficial in addressing missing data in PROMs and is now the recommended approach for dealing with missing data in PROMs (37).

Other challenges that are more PROM specific involve including the mixture of variables that are measured on different scales. Binary, categorical, ordinal and Likert scales are commonly used in PROMs. Comparing the relationship, content and scaling between PROMs is essential (20). Data preparation and statistical approaches are required that are appropriate for this type of variables and allow for analysis with variables on mixture of scales. Analysis of PROMs where there are repeated measurements over time creates yet another challenge and this has been raising much interest in the area of analysis of PROMs. Investigating trends in time and utilising the information on change in PROMs e.g. from the baseline or during or after treatment (38).

The preparation of PROMs data and choice of analytical techniques has to be carefully tailored to the scales of variables. Choosing between parametric and non-parametric statistics is particularly crucial. In a PROMs dataset when those questions are collected from a large number of participants, the variables (especially those within one domain) exhibit a high degree of collinearity that may pose a challenge for the analysis. Therefore, for the purpose of meaningful data analysis, pre-processing may require some degree of data reduction and/or variable selection. In some studies, the research question underpins justification for picking some symptoms and leaving others out for the PROM analysis. However, approaches that do not use any systematic variable selection will generally be criticised. For variable reduction some instrument specific scoring instructions suggest calculating average values for the items that belong to the same domain. However, this
indiscriminate approach can lead to a loss of important information. This is because not all the symptoms in the same domain are always highly correlated and thus can be condensed into one aggregate measure. To overcome this we recently proposed an approach that uses symptom clustering for variable reduction and selection. It was shown that clusters do not necessarily follow the domain pattern. We therefore recommend that average values, based on symptom clusters rather than domains, should be used (38).

Appraisal of the effect size and target of PROM instruments in radiotherapy

The main purpose of collecting PROMs is to improve patients' health related outcomes and increasingly in cancer survival studies focus is on the long term outcomes (39). Clinical utility of PROMS depends on the availability of information about clinical severity thresholds and what is important for a clinically meaningful change score for the patient (19). Kotronoulas and colleagues (12) found that effect sizes were small when evaluating PROM studies in cancer as interventions for improving outcomes and that the number of statistical findings within the evidence base is small. Two broad approaches to defining clinical effectiveness are recommended these include the use of distribution score information and the use of patient population rated “anchors” by which to estimate meaningful difference (19). A PROMS track record offers a practical indication of its psychometric properties so reviewing similar studies or patient populations can be a good way to choose a PROM. Widespread use does not mean that the tool has superior benefits but it is an advantage to compare across studies.

Conclusion

In conclusion can PROMs replace objective measures in radiotherapy? The answer is clearly yes but these tools require skill to use. The field of PRO measurement is a dynamic one and is moving fast, measures are becoming obsolete and new ones such as the US PROMIS generated by computer adaptive testing. Changes to workflow and increasing volume of patients mean that PROMS are a way to overcome shortcomings in toxicity reporting and can help clinical oncologists identify what is important for patients when they return to clinic and improve symptom management and communication. Including PROMs in routine practice will become more commonplace but research is needed to be able to best develop tools and content validity that is suitable for radiotherapy.

PROMs in research are well utilised but the quality of reporting is still poor in that large amounts of data is captured but not necessarily utilised in the final analysis. Developing the PRO research question and using the breadth of PROM concepts is essential to improve targeting of outcomes and generalizability of research. In the review of current studies few used symptom specific tools or patient generate index (PGI), which can mean that the condition specific tools may lack content validity for radiotherapy effects or individual problems that emerge as new treatments are developed. We need to consider the focus of the research question for survivorship research and the sensitivity of the PROM to be able to demonstrate change. There are limitations to PROMs in that the logistics of capture are time
consuming and require administrative skill, however new technologies are making this increasingly possible.

With the small number of radiation events future data pooling of PROMs and the ability to share between studies through data linkage promises exciting opportunities to study larger treatment effects and risk prediction. Do we need new PROM tools for radiotherapy or does it need to be driven by HRQOL groups and patients. PROMS are important for understanding patients experience and perceptions and therefore any new tools need to be embedded and follow the psychometric methodologies currently endorsed by the FDA. Patient generated tools are likely to be the future but as yet are not wide spread because of the technology requirements.

PROMS are an opportunity for clinical oncology because they offer the ability to assess patients' and collect long-term data often remotely. Changes in cancer care pathways means that patients receiving radiotherapy are less likely to be seen long term by their clinical oncologist (40) as changes in workflow and patient survivorship increase this means that follow up is less likely to be at the cancer centres. PROs provide benefits in demand management and can enhance services by providing innovation in co-ordinated process for surveillance, help patients make decisions and provide valuable observational data of population effects and differences in therapies that would not be seen in clinical trials.

Figure 1. Continuum from specificity to generalizability and the relationship of common cancer PROMs: the use of a range of tools provides a wider picture of patient impact.
Table 1: A sample of PROM instruments used in UK radiotherapy clinical trials from 2005-2014.

<table>
<thead>
<tr>
<th>Cancer site</th>
<th>Trial</th>
<th>PROM instrument</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bladder</td>
<td>BC2001</td>
<td>EORTC QLQ-C30</td>
<td>(41)</td>
</tr>
<tr>
<td>Brain</td>
<td>AVAGLIO</td>
<td>EORTC QLQ-C30, EORTC QLQ-BN20</td>
<td>(42)</td>
</tr>
<tr>
<td>Breast</td>
<td>START</td>
<td>EORTC QLQ-C30, EORTC QLQ-BR23, BIS, HADS</td>
<td>(43)</td>
</tr>
<tr>
<td>Head &amp; neck</td>
<td>PARSPORT</td>
<td>EORTC QLQ-C30, EORTC QLQ-H&amp;N35, Modified xerostomia questionnaire</td>
<td>(44)</td>
</tr>
<tr>
<td>Lung</td>
<td>SOCCAR</td>
<td>EORTC QLQ-C30, Euro-QoL EQ-5D</td>
<td>(45)</td>
</tr>
<tr>
<td>Prostate</td>
<td>RT01, CHHiP</td>
<td>SF-36, UCLA-PCI, FACT-P</td>
<td>(46) (47)</td>
</tr>
<tr>
<td>Rectum</td>
<td>CR07</td>
<td>SF-36, EORTC QLQ-CR38</td>
<td>(48)</td>
</tr>
<tr>
<td>Testicle</td>
<td>TE18</td>
<td>EORTC QLQ-C30, EORTC QLQ-TC26</td>
<td>(49)</td>
</tr>
</tbody>
</table>

Abbreviations: BIS: Body Image Scale; EORTC QLQ: European Organisation for the Research and Treatment Quality of Life Questionnaire, including C30 (Core 30), BN20 (Brain Cancer Module), BR23 (Breast Cancer Module), H&N35 (Head and Neck Module), CR38 (Colorectal Cancer Module), TC26 (Testicular Cancer Module); EQ-5D: Euro-QoL-5 Dimension; FACT-P: Functional Assessment of Cancer Therapy – Prostate; HADS: Hospital Anxiety and Depression Scale; SF-36: Short Form 36; UCLA-PCI: University of California Los Angeles Prostate Cancer Index.

References


