

Gibson A,¹ Longworth L,¹ Bennett B,² Shaw JW³

¹PHMR Ltd, London, UK; ²Bristol-Myers Squibb, Uxbridge, UK; ³Bristol-Myers Squibb, Lawrenceville NJ, USA

Background

- Cancer malignancies can differ in their pathology and presentation, although some symptoms (e.g., pain, unexplained weight loss, fatigue) are common across many cancer subtypes.
- Cancer symptoms and drug treatment effects can have a substantial burden on the lives of patients, therefore, capturing all relevant outcomes of people living with cancer is necessary to inform accurate decision-making for cancer technologies.
- The National Institute of Health and Care Excellence (NICE) currently recommend the use of the EQ-5D, a generic, preference-based measure (PBM), to assess patient reported outcomes (PROs) across interventions¹ and to allow for comparative assessments across unrelated health conditions and interventions.²
- However, generic PBMs may exclude concepts relevant to specific conditions, such as the three most frequently-reported cancer impacts of fatigue, difficulty sleeping, and worry.³
- Several concepts relevant to health-related quality of life (HRQoL) may also not be captured by existing condition-specific PBMs, including uncertainty, fulfilment of social roles, and separating cancer from other aspects of life.⁴
- Inclusion of missing concepts within future outcome measures could potentially allow for more holistic assessments of HRQoL.
- This study assessed the content validity of generic and condition-specific PBMs with persons treated for cancer, evaluated against ten COSMIN criteria for good content validity¹⁴, in order to best inform measurement strategies regarding the use of PBMs in oncology development programs as well as real-world applications.

Objectives

Objective 1

- To determine the content validity of five preference-based measures (PBM) and a set of newly developed, positively-oriented health items (PosiPRO), for measuring the outcomes of persons treated for cancer, using semi-structured interviews and questionnaire responses.

Objective 2

- To describe the relevancy, and any gaps in content validity, of the items and domains of five PBMs and the PosiPRO items for measuring the outcomes of persons treated for cancer, using semi-structured interviews.

Objective 3

- To describe the importance of the items within the five selected condition-specific and generic PBMs and the PosiPRO items for measuring the relevant outcomes of persons treated for cancer, using semi structured interviews.

Objective 4

- To compare the content validity of each of the five selected condition-specific and generic PBMs and the PosiPRO items in reporting the relevant outcomes of persons treated for cancer, using participants’ evaluations.

Methods

Design and conduct

- This non-interventional research study followed a qualitative design.
- Individual, semi-structured interviews were conducted with patients who were previously treated for cancer in the UK and the US.
- Interviews took place via teleconference with the aim of eliciting descriptions of the impact of cancer on their HRQoL and to assess the content validity of five PBMs and the novel PosiPRO items.
- The discussion guide that was developed followed guidelines for best practice in using semi-structured interviews to assess the content validity of existing questionnaire measures.⁵
- The study design, materials, and patient-facing documents received appropriate ethical approval by an Institutional Review Board.
- Interviews were recorded, transcribed, pseudonymized, and checked for accuracy throughout the data collection process, with the option to retrain interviewers where required.
- Transcripts were analysed using thematic and content analytical approaches in accordance with best practice guidelines for qualitative interviews in health economics.^{6,7}

Questionnaires

- Participants were each asked to complete and assess the content validity of only two questionnaires within their interview, to minimise the risk of cognitive overload for each participant. Assessed questionnaires included five PBMs and one newly developed set of positively framed health items.

- Each participant received copies of the two questionnaires by mail in advance of the interviews. Questionnaires were administered in English.

Generic PBMs:

- EQ-5D-5L⁸
- EQ-HWB (short form)⁹
- CS-Base¹⁰

Cancer-specific PBMs

- EORTC QLU-C10D¹¹
- FACIT-8D¹²

Newly developed items:

- PosiPRO items. Twelve generic HRQoL items to measure positive aspects of health due to limitations in existing instruments that focus on the presence of disease and their negative impacts (Pickard S, 2021 - personal communication)

- Participants were sequentially allocated into six groups, containing those who have recently or are currently receiving treatment for cancer as well as those who were considered long-term survivors of cancer (Figure 1).

Figure 1. Participant questionnaire group allocation

Group 1	Group 2	Group 3
PosiPRO items	QLU-C10D	EQ-5D-5L
FACIT-8D	EQ-5D-5L	PosiPRO items
Group 4	Group 5	Group 6
EQ-HWB-S	FACIT-8D	CS-Base
QLU-C10D	CS-Base	EQ-HWB-S

Sample

- Ninety-six persons treated for cancer were interviewed, inclusive of a subsample from the UK (n=47, 49%) and US (n=49, 51%) (Table 1, Table 2).
- Patients were required to self-report a current haematological or solid tumour cancer diagnosis from a medical practitioner (Figure 2).
- Participants were required to be a current or recent recipient (within the past year) of drug treatment for cancer or have received a first drug treatment more than two years ago but not more than 10 years ago.
- Patients were not eligible for participation if they were currently undergoing treatment as part of a clinical trial or had received curative therapy and considered themselves to be cancer free.
- Sixty-nine patients reported a comorbid condition (72%), including high blood pressure (n=18, 19%), high cholesterol (n=9, 9%), diabetes (n=8, 8%).

- A sampling limitation is that the study included a large proportion of patients with breast cancer (n=51, 53%), however the remainder of the sample contained a wide range of initial diagnosis types. Evidence also suggests that the relative burden of cancer may be comparable across diagnoses.¹³

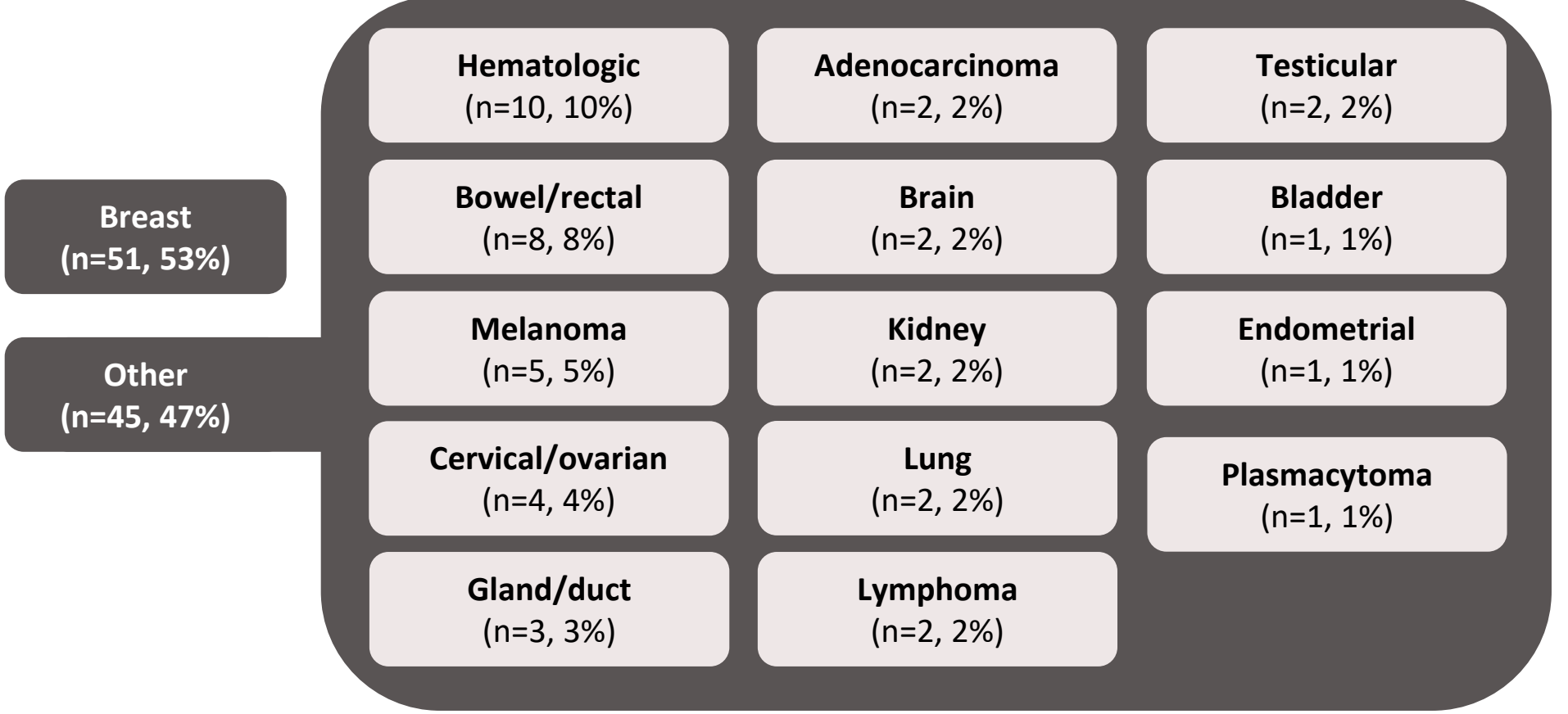
Table 1. Patient demographic characteristics

(n, %)	
Country	
United Kingdom	47 (49)
USA	49 (51)
Age (years)	
Mean (SD)	53 (11)
Range	28-76
Study sub-group	
Current/recently treated	46 (48)
Long-term survivor	50 (52)
Occupation	
Retired	19 (20)
Unemployed	17 (18)
Voluntary	2 (2)
Employed	56 (58)

Table 2. Patient clinical characteristics by subgroup

Diagnosis	All (N=96)	Current/recent (n=46)	Long-term survivor (n=50)
Age at diagnosis (M,SD)	46 (11)	47 (11)	45 (12)
Multiple cancer diagnoses (n, %)	42 (44)	28 (61)	14 (49)
No evidence of disease progression (n, %)	7	-	7

Figure 2. Initial cancer diagnosis self-reported during interview



Findings

- Questionnaires were evaluated against ten criteria across three COSMIN domains for good content validity: relevance, comprehensiveness, and comprehensibility.¹⁴
 - These three domains should be considered separately when evaluating content validity.

Objective 1: Content Validity by Questionnaire

- For appraisals of the questionnaires as a whole, there was no evidence that the EQ-HWB-S, QLU-C10D, and FACIT-8D, overall, might not adequately capture the impact of cancer or treatment side effects:

EQ-HWB-S

‘I think it’s a good general questionnaire. In my situation right now, I feel like it covers all the bases, and I am able to answer it’. (US168, current, breast)

QLU-C10D

‘It depends where you are in your journey... So there’s those things that impact me every day. But the things like the constipated or nausea, that was the stuff that I had when I was in throes of the treatment’. (UK726, long-term, breast)

FACT-8D

‘When you’re a cancer patient, you’re often just seen as your medical condition, and nobody really has the time to speak to you about yourself, as a whole person. ... This, at least, addresses those sorts of aspects, which you [aren’t] often asked about, when you go for your medical appointment’. (UK25, current, breast)

- There was at least some evidence of potential difficulties or inappropriateness, overall, for cancer outcome measurement for the EQ-5D-5L, CS-Base, and PosiPRO items.
- All individual items for the EQ-5D-5L, EQ-HWB-S, QLU-C10D, and FACIT-8D, were each evaluated as relevant for the context of use, with patients providing reasons for their responses directly linked to their symptom and treatment experiences, or directly stating their appropriateness:

EQ-5D-5L

‘because that’s when you know that you’re not well, if you can’t wash or dress yourself’. (UK726, long-term, breast) – Self-care

QLU-C10D

‘Sometimes I just don’t feel hungry at all. So, it’s a good question for people going through treatment’. (US270, current, breast) - Appetite

Objective 2: Measurement gaps

- All questionnaires performed similarly on the COSMIN dimension of comprehensiveness, with all reported as having missing aspects (Table 3).

Table 3. Missing aspects by questionnaire

Missing aspect	Questionnaire count				Total participants (n, %)
	Total	Generic	Condition-specific	PosiPRO	
Healthcare and physical support needs	6	3	2	1	26 (27)
Appetite/eating/diet	6	3	2	1	17 (18)
Relationships and family	5	3	1	1	17 (18)
Emotional well-being in addition to depression	4	1	2	1	28 (29)
Financial impacts	4	2	1	1	5 (5)
Longer-term impacts of cancer and treatment	3	2	0	1	12 (13)

- There lacked consensus on the preferred timeframe for measurement of cancer outcomes. Reasoning suggested this was due to the variation in symptoms and impacts across the cancer journey, a given day, or a given week:

“Cancer can tend to be a roller coaster. And how you are feeling today, how you are feeling the past seven days is not necessarily indicative of how you have been feeling generally over the past month or two months... Cancer is a journey. And looking at it as a one-day or even a seven-day, I don’t think gets to what the state of my health or well-being is with living with cancer. I think it’s a false premise”. (US169, long-term, lymphoma)

Objective 3: Importance of included items

- Each individual item for each questionnaire was compared to symptoms and impacts reported spontaneously by patients before questionnaire completion.
- Most items across questionnaires were reported by patients as symptoms or impacts and did not receive negative appraisals from patients regarding relevancy for the target population. This was true for all EQ-HWB-S, QLU-C10D, and FACIT-8D items.
- The concepts of vision and hearing (CS-BASE) were not commonly reported as symptoms or impacts by most patients nor individually appraised by patients as relevant. One patient with astrocytoma did report some vision impact.

Objective 4: Content validity comparisons across questionnaires

- All questionnaires were evaluated in terms of criteria contributing to the three COSMIN domains of relevance, comprehensiveness, and comprehensibility (Figure 3). Interview questions, symptoms and impacts, and qualitative groupings of responses were mapped to the 10 criteria.
- Relevancy of the FACIT-8D, EQ-HWB-S, and QLU-C10D were comparable, differing in terms of the response options and recall periods only, rather than on the content of the items.

Figure 3. Questionnaire evaluations in terms of the COSMIN criteria

COSMIN Criteria	No evidence of problems (no completing participants)	Some evidence of problems	Substantial evidence of problems (>25% completing participants)
Relevance			
Are the included items relevant for the construct of interest?	EQ-HWB-S QLU-C10D FACIT-8D	EQ-5D-5L CS-Base PosiPRO	
Are the included items relevant for the target population of interest?	EQ-HWB-S QLU-C10D FACIT-8D	EQ-5D-5L CS-Base PosiPRO	
Are the included items relevant for the context of use of interest?	EQ-5D-5L EQ-HWB-S QLU-C10D FACIT-8D	CS-Base PosiPRO	
Are the response options appropriate?	EQ-5D-5L FACIT-8D	EQ-HWB-S QLU-C10D PosiPRO	CS-Base
Is the recall period appropriate?	EQ-HWB-S CS-Base FACIT-8D	QLU-C10D PosiPRO	EQ-5D-5L
Comprehensiveness			
Are no key concepts missing?		EQ-5D-5L EQ-HWB-S CS-Base QLU-C10D FACIT-8D PosiPRO	
Comprehensibility			
Are the PROM instructions understood by the population of interest as intended?	FACIT-8D	EQ-5D-5L EQ-HWB-S CS-Base QLU-C10D PosiPRO	
Are the PROM items and response options understood by the population of interest as intended?	EQ-5D-5L	EQ-HWB-S CS-Base QLU-C10D FACIT-8D PosiPRO	
Are the PROM items appropriately worded?			EQ-5D-5L EQ-HWB-S CS-Base QLU-C10D FACIT-8D PosiPRO
Do the response options match the question (e.g., symptom better queried as intensity or frequency, dichotomous response options)	EQ-5D-5L	EQ-HWB-S QLU-C10D	CS-Base FACIT-8D PosiPRO

Conclusions

- All measures were generally seen by patients as capturing appropriate aspects of HRQoL for cancer outcomes.
- Together, condition-specific measures were evaluated as having better relevancy than generic PBMs. The FACIT-8D was evaluated as having better content validity in terms of relevancy, although the EQ-HWB-S received the most favorable evaluation of relevancy for generic PBMs.
- All measures demonstrated comparable comprehensiveness, with all suggested to be missing concepts considered by patients as important to measure for their cancer and cancer treatment experiences.
- The EQ-5D-5L was evaluated best in terms of comprehensibility. This was followed by the QLU-C10D and FACIT-8D; both received similar evaluations.
- Language and recall period were frequently challenged. Patients suggested language and format changes to multiple items across all measures. Recall periods were frequently raised as problematic due to the substantial variation in cancer symptoms within a given day or week.
- The PosiPRO items saw less favorable evaluations of content validity than other measures but the positive evaluations from participants suggested they may have the potential for filling gaps in cancer outcomes measurement.
- Further HRQoL instrument development is encouraged. Bolt-on items to the existing positively appraised questionnaires should be considered with regards to measurement gaps (e.g., appetite, eating or diet, and long-term impacts of cancer and treatment). A new cancer-specific tool could be considered, inclusive of conventional health items, treatment impacts, psychological items, and positively oriented items. Such a development can be halfway between a bolt-on and new measure with a large potential for impact given the high proportions of HTA submissions within oncology.

References

1. NICE. The reference case 2013 [Available from: <https://www.nice.org.uk/process/pmg9/chapter/the-reference-case>.
2. Brazier et al. A Review of Generic Preference-Based Measures for Use in Cost-Effectiveness Models. Pharmacoeconomics. 2017;35(1):21-31.
3. Stark et al. The Symptom Experience of Patients with Cancer. Journal of hospice and palliative nursing : JHPN : the official journal of the Hospice and Palliative Nurses Association. 2012;14(1):61-70
4. Park et al. The value of immunotherapy for survivors of stage IV non-small cell lung cancer: patient perspectives on quality of life. J Cancer Surviv. 2020;14(3):363-76.
5. Brod et al. Qualitative research and content validity: Developing best practices based on science and experience. Qual Life Res. 2009;18:1263-1278
6. Clarke & Braun. Thematic analysis. In: Michalos AC, editor. Encyclopaedia of Quality of Life and Well-Being Research. Dordrecht, Netherlands: Springer; 2014.
7. Coast & Jackson. Understanding primary data analysis. In: COAST J, editor. Qualitative methods for health economics: Rowman & Littlefield.; 2017.
8. Herdman et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
9. Brazier et al. The EQ-HWB: Overview of the development of a measure of health and wellbeing and key results. Value in Health. 2022;25(4):482-91.
10. Krabbe et al. Patient-Centered Item Selection for a New Preference-Based Generic Health Status Instrument: CS-Base. Value in Health. 2019;22(4):467-73.
11. King et al. QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLU-C30. Qual Life Res. 2016;25(3):625-36.
12. King et al. The Functional Assessment of Cancer Therapy Eight Dimension (FACT-8D), a Multi-Attribute Utility Instrument Derived From the Cancer-Specific FACT-General (FACT-G) Quality of Life Questionnaire: Development and Australian Value Set. Value in Health. 2021;24(6):862-873
13. Pickard, SA., Jiang, R., Lin, HW., Rosenbloom, S., & Cella, D. Using patient-reported outcomes to compare relative burden of cancer: EQ-5D and functional assessment of cancer therapy-general in eleven types of cancer. Clinical Therapeutics. 2016;38(4):769-777.
14. Terwee et al. COSMIN methodology for assessing the content validity of PROMs. : COSMIN: Consensus-based standards for the selection of health measurement instruments; 2018.

Acknowledgments

- This research was funded by Bristol Myers Squibb
- This study would not have been possible without the patients who were willing to take part in research interviews