VALUE IN HEALTH 15 (2012) 305–311

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Patient Reported Outcomes

Comparison of FACT- and EQ-5D–Based Utility Scores in Cancer

A. Simon Pickard, PhD[1,], Saurabh Ray, PhD[2], Arijit Ganguli, PhD[2], David Cella, PhD[3]*

[1]*Second City Outcomes Research, Oak Park, IL, USA;* [2]*Global Health Economics and Outcomes Research, Abbott Laboratories, Abbott Park, IL, USA;* [3]*Northwestern University, Feinberg School of Medicine, Chicago, IL, USA*

A B S T R A C T

Objective: Although utility-based algorithms have been developed for the Functional Assessment of Cancer Therapy (FACT), their properties are not well known compared with those of generic utility measures such as the EQ-5D. Our objective was to compare EQ-5D and FACT preference-based scores in cancer patients. **Methods:** A retrospective analysis was conducted on cross-sectional data collected from 472 cancer patients who completed both FACT-General and the EQ-5D. Preference-based scores were calculated by using published scoring functions for the EQ-5D (Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095–108; Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005;43:203–20) and FACT (Dobrez D, Cella D, Pickard AS, et al. Estimation of patient preference-based utility weights from the Functional Assessment of Cancer Therapy-General. Value Health 2007;10:266–72; Kind P, Macran S. Eliciting social preference weights for Functional Assessment of Cancer Therapy-Lung health states. Pharmacoeconomics 2005;23:1143–53; Cheung YB, Thumboo J, Gao F, et al. Mapping the English and Chinese versions of the Functional Assessment of Cancer Therapy-General to the EQ-5D utility index. Value Health 2009;12:371–6). Scores were compared on the basis of clinical severity by using Eastern Cooperative Oncology Group performance status ratings by physicians and patients. Relative efficiency of each scoring function was examined by using ratios of F statistics. **Results:** Mean scores for the overall cohort were lowest when using Kind and Macran's FACT UK societal algorithm (0.55, SD 0.09) and highest when using Dobrez et al.'s FACT US patient algorithm (0.83, SD 0.08). Mean difference scores associated with clinical severity, when extrapolated to quality-adjusted life-years (QALYs), had a range of 0.18 QALYs gained using FACT (Kind and Macran) to 0.45 QALYs gained using the EQ-5D (Dolan). However, relative efficiencies suggested that FACT (Kind and Macran) scores may provide greater statistical power to detect significant differences based on clinical severity. **Conclusions:** We found important differences in utilities scores estimated by each algorithm, with FACT-based algorithms tending to underestimate the QALY benefit compared with algorithms based on the EQ-5D. These differences highlight some of the challenges in using disease-specific preference-based measures for decision making despite potentially more relevant disease-specific content.

Keywords: cancer, EQ-5D, health-related quality of life, health state utilities, utility assessment.

Introduction

The ability to generate utility scores as an outcome in oncology trials is essential to the conduct of cost-utility analyses. Utility scores enable the calculation of quality-adjusted life-years (QALYs) [1], a metric that adjusts time in a health state by the desirability or preference for that health state to evaluate the value and/or cost-effectiveness of treatments for cancer. The most widely used utility measures are generic measures, particularly the EQ-5D [2]. Numerous national value sets are available to score the EQ-5D, with the most highly cited societal value sets being derived from general population in the United Kingdom [3] and the United States [4]. The EQ-5D has generally demonstrated validity and reliability in studies of cancer [5]. There are often instances, however, when utility scores are desired but no preference-based measure is used in a study, or the generic measure may lack responsiveness to meaningful changes in health-related quality of life. The former issue has fostered research focused on the mapping of non–preference-based disease-specific measures to generic measures that generate utility scores [6].

Preference-based scoring functions for cancer-specific measures include several for the Functional Assessment of Cancer Therapy (FACT) [7], a well-established family of cancer-specific measures. Originally developed by using methods based on psychometric theory, the FACT scoring involves the summation of ordinal-level responses to items belonging to each scale. This approach contrasts with the "utility" approach to health measurement, where a summary score is derived by applying a utility function or set of preference weights assigned to the levels and dimensions of the measure.

The preference-based algorithms published for the FACT system have varied in their methods and rater perspective. Dobrez

* *Address correspondence to:* A. Simon Pickard, Second City Outcomes Research, 414 N. Ridgeland Avenue, Oak Park, IL 60302, USA.
E-mail: simon.pickard@secondcityoutcomesresearch.com.

doi:10.1016/j.jval.2011.11.029

Table 1 – Utility-based algorithms.

Author, year [reference]	Measure	Domains in measure	Items/domains in algorithm	Source of utilities	Range of possible values
Dolan, 1997 [3]	EQ-5D (3-level)	Mobility, usual activities, self-care, pain/discomfort, anxiety/depression	Mobility, usual activities, self-care, pain/discomfort, anxiety/depression	Societal (UK)	−0.59 to 1.0
Shaw et al., 2005 [4]	EQ-5D (3-level)	Mobility, usual activities, self-care, pain/discomfort, anxiety/depression	Mobility, usual activities, self-care, pain/discomfort, anxiety/depression	Societal (US)	−0.109 to 1.0
Dobrez et al., 2007 [8]	FACT-G	Physical well-being (PWB), emotional well-being (EWB), functional well-being (FWB), social well-being (SWB)	Two items: PWB, Two items: FWB	Cancer patients (US)	0.50–1.04
Kind and Macran, 2005 [9]	FACT – Lung (FACT-L, version 4)	PWB, EWB, FWB, SWB, Lung component – symptoms (LCS)	PWB, EWB, FWB, SWB, LCS	Societal (UK)	0.18–0.70
Cheung et al., 2009 [11]	FACT-G (version 4)	PWB, EWB, FWB, SWB	Scale scores for PWB, EWB, FWB	Mapped from Dolan 1997 (UK societal)	0.238–0.982

FACT-G, Functional Assessment of Cancer Therapy-General.

et al. [8] estimated a set of value weights for selected items from FACT-General based on time trade-off (TTO) scores directly elicited for own health from US cancer patients. Kind and Macran [9] derived a set of societal preference weights by directly eliciting visual analogue scale (VAS) ratings from the general population in the United Kingdom for a descriptive health classifier system based on FACT-Lung (FACT-L). Cheung and colleagues mapped both English and Chinese versions of FACT to the EQ-5D utility scores and derived a single mapping function for both languages based on three of the four summary scores from FACT.

Given the many differences between the EQ-5D and FACT-based scoring functions, we would expect them to generate different scores. The extent to which the scoring functions differ, however, has not been well documented or contrasted. Because different decision makers have different needs, they may wish to not only consider the value that each approach can contribute to decision making but also understand how the metrics differ relative to each other, particularly with respect to statistical efficiency and QALY calculations. Thus, our objective was to examine preference-based scores generated by EQ-5D and FACT scoring functions to better understand the strengths and limitations of each approach in valuing health.

Methods

Data

A retrospective analysis was conducted on a data set where patients completed both the EQ-5D and FACT. The cohort consisted of 534 cancer patients who participated in a US-based multicenter symptom scale validation study, which has been previously described [10]. Patients completed both instruments on the same day. All patients had advanced cancer classified as relating to any 1 of 11 tumor sites: bladder, brain, breast, colorectal, head/neck, hepatobiliary/pancreas, kidney, lung, lymphoma, ovary, or prostate. Approximately equal proportions of male and female patients were recruited for the non–gender-specific cancers.

Patients were recruited from six sites that were geographically representative of National Comprehensive Cancer Network member institutions, an alliance of National Cancer Institute–approved comprehensive cancer centers: Duke University Medical Center, Fred Hutchinson Cancer Research Center, Dana Farber Cancer Institute, H. Lee Moffitt Cancer Center and Research Institute, and the Robert H. Lurie Comprehensive Cancer Center at Northwestern University. In addition, patients were recruited by members of the Cancer Health Alliance of Metropolitan Chicago, a coalition of four community support agencies serving the Chicago metropolitan area. The Cancer Health Alliance of Metropolitan Chicago organizations provide social, emotional, and informational support services to cancer patients free of charge and are unaffiliated with a medical center or university. Each Cancer Health Alliance of Metropolitan Chicago agency serves different geographical and sociodemographic cancer patient populations.

Measures/algorithms

We compared the scores generated by five algorithms/scoring functions, two of them based on the EQ-5D and three based on responses to FACT (Table 1). They are described in greater detail below.

EQ-5D

The EQ-5D descriptive system consists of five dimensions (Mobility, Self-Care, Usual Activities, Pain/Discomfort, and Anxiety/Depression), each with three levels (3L) of health [2]. The EQ-5D also includes a 20-cm VAS, which asks the respondents to rate their health today from 0 (worst imaginable health) to 100 (best imaginable health). A preference-based score is calculated from responses to the health state descriptive system that is typically interpreted along a continuum where 1 represents best possible health and 0 represents dead, with some health states being worse than dead (<0). Participants were asked to complete the standard US English language EQ-5D (3L) and the VAS. The US and UK English language versions of the EQ-5D are nearly identical, differing only in the instruction to place a tick (UK) or checkmark (US) in the box that best describes your own health state today.

The EQ-5D preference-based scores were calculated by using the algorithms developed by Dolan [3] from the general population in the United Kingdom and by Shaw and colleagues [4] for the United States. A ceiling effect is observed in milder health conditions with the EQ-5D health classifier system [10,12], a limitation that applies equally to both algorithms. The scores generated by the Dolan UK value set range from −0.59 (for health state vector 33333, which represents the worst health state) to 1.0 (for health state vector 11111, which represents full health), while the scores for the US value set from Shaw and colleagues cover a smaller range of scale, from −0.109 to 1.0.