

Comparison of SF-6D and EQ-5D Scores in Patients With Breast Cancer

Mahmood Yousefi,¹ Safa Najafi,² Shahram Ghaffari,³ Alireza Mahboub-Ahari,¹ and Hossein Ghaderi^{4,*}

¹Iranian Center of Excellence in Health Management, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, IR Iran

²Iranian Center for Breast Cancer (ICBC), Tehran, IR Iran

³Social Security Organization (SSO), Tehran, IR Iran

⁴Health Economics Department, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, IR Iran

*Corresponding Author: Hossein Ghaderi, Health Economics Department, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, IR Iran. Tel: +98-2188794302, Fax: +98-2188883334, E-mail: hoss_ghaderi@yahoo.com

Received 2014 September 11; Revised 2014 September 28; Accepted 2014 October 22.

Abstract

Background: Utility values are a key component of a cost-utility analysis. The EQ-5D and SF-6D are two commonly used measures for deriving utilities. Of particular importance is assessing the performance of these instruments in terms of validity.

Objectives: This study aimed to compare the performance of the EQ-5D and the SF-6D in different states of breast cancer.

Patients and Methods: This was a cross-sectional study of 163 patients with breast cancer who attended the breast cancer subspecialty clinic affiliated with the breast cancer research center (BCRC) at ACECR, in Tehran, Iran, and were consecutively recruited. Patients completed several questionnaires, including the EQ-5D, SF-36, and general questions regarding their demographic characteristics. Utility values for different states of breast cancer were obtained using predetermined algorithms for the EQ-5D and SF-6D. The distribution of the utility values and the differences between the different states for both instruments were statistically assessed. Furthermore, the agreement between the two instruments was evaluated using intra-class correlation coefficients and Bland-Altman plots.

Results: The mean and median EQ-5D utility scores for the total sample were 0.685 and 0.761, respectively. The mean SF-6D utility score for the total sample was 0.653, and the median utility score was 0.640. The mean utility values of the EQ-5D for "state P," "state R," "state S," and "state M" were estimated as 0.674, 0.718, 0.730, and 0.552, respectively. The SF-6D provided mean utility values of 0.638, 0.677, 0.681, and 0.587 for those states. Both instruments assigned statistically significant ($P < 0.01$) scores for different states. The intra-class correlation for the two measures was 0.677 (95% confidence interval (CI): 0.558 - 0.764). The Bland-Altman plot indicated a better agreement on the higher values and that at higher values, the EQ-5D yields a higher score than the SF-6D; this relationship was reversed at lower values.

Conclusions: Although the two instruments were able to discriminate between various states, the values derived from these instruments were quite different. This distinction could have influenced the conclusions of an economic evaluation. Further research is required to determine which instrument should be used in economic evaluations.

Keywords: Quality of Life, Breast Cancer, Utilities, EQ-5D, SF-6D

1. Background

Cost-utility analysis as one of the most frequently used methods of economic evaluation, and it uses the Quality Adjusted Life Years (QALYs) as a generic measure of health outcome. QALYs are unique and exact measures that potentially permit comparisons both within and across all interventions (1). QALYs capture both the quantity and quality of life years in a single measure of health outcome. The quality component of the QALYs is measured in terms of utility. Utility is a measure to reveal preferences for a given health state that range from 0 (death) to 1 (full health), although a negative value (states perceived to be worse than death) is also possible (2). There are multiple generic preference-based instruments used to estimate utility values for computing QALYs. The EuroQol (EQ-5D) (3-5) and Short Form (SF-6D) (6, 7) are two of the most widely used instruments. Both instruments use a spe-

cific descriptive system to classify different health states. Based on the scoring algorithm, which is derived from the general public specifically for each instrument, each classified health state is assigned a value (8). However, previous studies that have compared these instruments have revealed some differences in their performances (9). Even though the EQ-5D is widely used, easily administered, and able to detect large differences in health status (10), its poor performance in detecting small changes in high level of utilities can be of concern (11). In order to enhance the descriptive richness and sensitivity, the SF-6D measure was therefore developed from the most widely used multi-dimensional SF-36 instrument (7, 12). The main purpose behind its development was to incorporate both the descriptive richness and preference-based properties into a single measure (6). Meanwhile,

previous studies indicated that the SF-6D over-predicted at lower levels of the utility values (6) and revealed a low sensitivity to change within a lower range of utilities (13). Moreover, evidence suggests that these two instruments report different utilities and levels of agreements for similar clinical conditions (8, 14-16). These facts highlight the necessity of more investigations and exercising caution when using these instruments for decision-making.

Breast cancer is the leading cause of cancer death among women and the most frequently diagnosed cancer worldwide (17). For chronic diseases, such as breast cancer, which have a demonstrated impact on health-related quality of life (HRQL) (18-20), integrating HRQL data into the treatment evaluation is essential. Several studies have compared these two instruments in diseases other than breast cancer (8, 9, 11, 13, 16). To the best of our knowledge, no studies have thus far compared the performance of these instruments in Iranian patients with breast cancer, assessing the performance of these instruments in different diseases and sociocultural contexts is of great importance.

2. Objectives

This study aimed to compare the performance of EQ-5D and SF-6D in different states of breast cancer.

3. Patients and Methods

3.1. Design and Recruitment

This was a cross-sectional study that was approved jointly by the ethics committees of Iran University of Medical Sciences (IUMS) and breast cancer research center (BCRC), ACECR in Tehran, Iran, in December 2012 with code number 631. A total of 163 patients with breast cancer who attended a breast cancer subspecialty clinic affiliated with the BCRC was consecutively recruited between November 2013 and June 2014. The sufficiency of the sample size was based on the minimum value for the intraclass correlation coefficient (ICC) to be treated as good agreement (21) and also on the method proposed by Walter for determining the sample size for ICC (with $\alpha = 0.05$ and $\beta = 0.20$) (22). The criteria for inclusion of patients in the study were as follows: confirmed pathological breast cancer providing that the clinical history of the patient was registered in the clinic's database, the patient provided written informed consent, and that there was no comorbidity. Literate patients were asked to self-administer the questionnaire package while they were in the waiting area of the clinic for a physician's visit. This package included the EQ-5D, the SF-36, and questions regarding their demographic characteristics. For illiterate patients, the questionnaire was administered by a trained research assistant. The total sample consisted of patients with different stages of the disease, including the early stages from I - III (both primary breast cancer and loco-regional

recurrence) and metastasis. Due to the insufficiency of the data, contralateral breast cancer patients were not included in this study. Patients were assigned to different states based on the date of diagnosis and their pathological stages as recorded in the database. Since the obtained utilities were assumed to be used in economic modeling, the states were constructed based on the predefined states by Lidgren et al. (23). Accordingly, the first year after primary breast cancer and that after recurrence are defined as "State P" and "State R," respectively. The second and following years after primary breast cancer or recurrence are defined as "State S," while metastatic disease is termed "State M." Further details about the definitions of these states can be found elsewhere (23).

3.2. Instruments

EQ-5D: the EQ-5D (3, 5) is a standard and generic preference-based instrument that was developed by a European group for driving utility values. It has a descriptive classification system with five domains (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). Each domain has three possible responses (no problems, some problems, and extreme problems). The descriptive system of the EQ-5D offers a total of 243 health states. Originally, these health states were assigned values, using the time trade-off method from a sample of 3395 respondents of the United Kingdom's (UK) general population. Several tariffs from different countries are available. For this study, the UK scoring algorithm was applied. This algorithm yields scores ranging from -0.59 to 1.00 , with 0 representing being dead, 1.00 indicating a state of full health, and negative scores indicating health states worse than being dead. The reliability and validity of the EQ-5D have been well documented in different contexts for different diseases (3, 11, 13).

SF-6D: the SF-6D (6), a multidimensional health classification system, is derived from the SF-36, a generic health status instrument that consists of eight scales (12, 24). The SF-6D uses 11 questions from the SF-36 to define the six domains (physical functioning, role limitation, social functioning, pain, mental health, and vitality). Each domain has between four and six levels, resulting in a total of 18000 health states. The health states for SF-6D are assigned values using the algorithm produced by Brazier and colleagues (6, 7). A sample of 249 health states using the standard gamble (SG) method were applied to derive preference weights from a representative sample of the UK population. The utility scores obtained from the SF-6D ranged from 0.29 to 1.00 , with 0.29 indicating the worst health state and 1.00 suggesting a fully healthy state other than pet toys. The SF-6D has demonstrated good reliability and validity in different settings (11).

3.3. Statistical Analysis

Descriptive statistics (mean, median, inter-quartile

range (IQR), and standard deviation (SD)) were used to characterize the study sample. The normality of the utility scores was tested graphically and with the skewness and kurtosis normality test. Due to a skewed distribution of the data, we performed a nonparametric Kruskal-Wallis test to test for differences in utility values between different states for both EQ-5D and SF-6D. In addition, we pooled the entire sample of different states together and assessed the agreement between two instruments by the ICC and Bland-Altman plot. All statistical analyses were performed using STATA 12.0 and MedCal 13.0.6 software.

4. Results

4.1. Patient Characteristics

Of the 163 patients who were eligible to participate in the study, two (1.2%) failed to complete the EQ-5D (step 5) and three (1.8%) did not complete the SF-6D. We assessed the characteristics of the failed patients and found that they did not differ from the patients who fully completed the instruments. Therefore, these missing data were supposed to be at random. We performed our analyses based on the 158 patients for whom both the EQ-5D and SF-6D were available. The clinical and demographic characteristics of patients are presented in Table 1. The mean age of respondents was 46.7 years. The largest and smallest numbers of patients were in the "state S" (44.9%) and "state R" (9.4%) groups, respectively.

4.2. Instrument Comparisons

As illustrated in Table 2, the mean and median EQ-5D utility scores for total sample were 0.685 and 0.761, respectively. At the same time, the mean SF-6D utility score for the total sample was 0.653, and the median utility score was 0.640. The mean utility scores for different states of the EQ-5D for "state P," "state R," "state S," and "state M," were estimated at 0.674, 0.718, 0.730, and 0.552, respectively. The SF-6D gave the mean utility values of 0.638, 0.677, 0.681, and 0.587 for those states. Moreover, the median for different states and the IQR for the total sample are shown in Table 2. Figure 1 and 2 show the distribution of utility values in the total sample for both EQ-5D and SF-6D measures. It is immediately obvious that the distribution of utility values for the SF-6D was more symmetrical than that for EQ-5D. In addition, compared to SF-6D, the range of utility values for EQ-5D was widely spread. Furthermore, the skewness and kurtosis normality test revealed a skewed distribution ($P = 0.001$) for EQ-5D, whereas the utility scores for SF-6D were normal ($P = 0.115$). Both EQ-5D (K-W, $P = 0.004$) and SF-6D (K-W, $P = 0.002$) gave significantly different utility scores among different health states. The ICC for two measures was 0.677 (95% CI: 0.558 - 0.764). The Bland-Altman plot used for agreement is presented in Figure 3. In general, the pattern of observations indicates that there is better agreement amongst the higher values, and that at higher values, the EQ-5D yields a higher score than the SF-6D, while this relationship reverses at lower values.

Table 1. Clinical and Demographic Characteristics of Patients with Breast Cancer^a

Characteristics	Values
Entire states	158 (100)
State P	48 (30.3)
State R	15 (9.4)
State S	71 (44.9)
State M	24 (15.1)
Mean age, y	46.7 (9.97)

^aData are represented as No (%), except for age that is represented by mean (SD).

Table 2. Utility Scores of Instruments

	SF-6D				EQ-5D			
	Median	Mean	SD	IQR	Median	Mean	SD	IQR
Total sample	0.640	0.653	(0.129)	0.120	0.761	0.685	(0.216)	0.327
State P	0.638	0.638	(0.125)	-	0.761	0.674	(0.201)	-
State R	0.680	0.677	(0.063)	-	0.761	0.718	(0.139)	-
State S	0.671	0.681	(0.134)	-	0.799	0.730	(0.221)	-
State M	0.581	0.587	(0.130)	-	0.511	0.552	(0.227)	-

Abbreviation: IQR, Inter-quartile range.

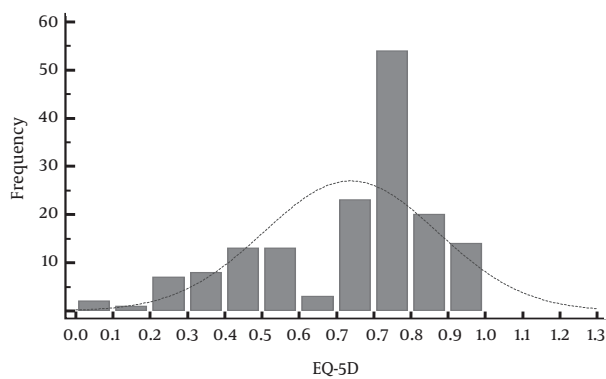


Figure 1. Distribution of the EuroQol (EQ-5D) Utility Scores

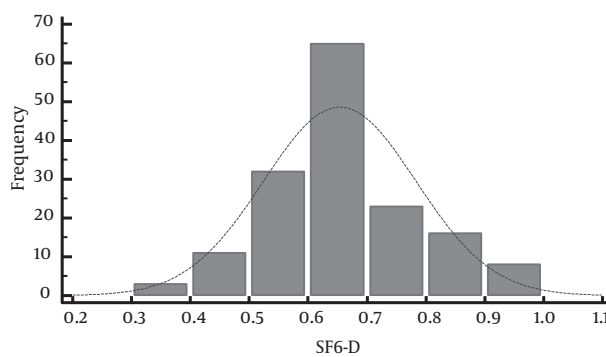


Figure 2. Distribution of the Short Form 6D (SF-6D) Utility Scores

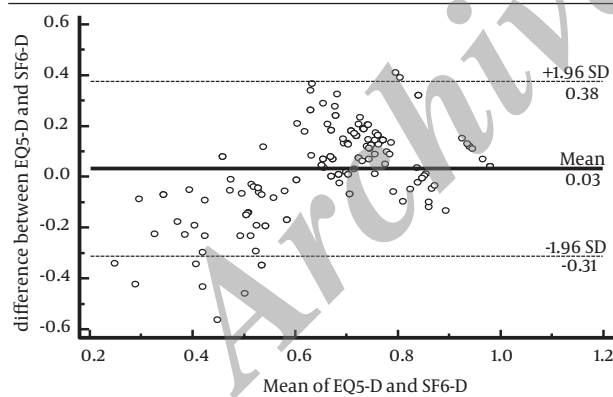


Figure 3. Bland-Altman Plots of Agreement between EQ-5D and SF-6D Utilities

5. Discussion

Utility indices are the key component of a cost-utility analysis. The instruments which are used to produce utilities should be valid, and using different instruments should not influence the conclusions of an economic evaluation. Therefore, assessing the performance of different utility measures across various interventions and sociocultural contexts is of great importance. We compared the performance of the EQ-5D and the SF-6D in patients

with different states of breast cancer. In general, the two instruments were able to discriminate between various states, and they assigned low values to most severe states and also conversely to states with a low severity. Despite the fact that these instruments were primarily developed to measure exactly similar values, we found considerable differences between these instruments. Firstly, the EQ-5D and SF-6D captured different numbers of domains (5 vs. 6), and the vitality domain within the SF-6D has no counterpart domain within the EQ-5D. Secondly, the two instruments revealed considerably different distributions (Figures 1 and 2). The EQ-5D was distributed with skewing towards the higher values, which produced a ceiling effect. In contrast, the SF-6D distributed with an approximation to a normal distribution. Due to dissimilar distributions, we compared the median for two measures and, consequently, the differences between the median values exceeded the minimally important difference (MID) (25) for both instruments. Even though the ICC showed a fair agreement, the Bland-Altman plot represented that a large proportion of differences were exceeded the MID. Furthermore, the agreement was much poorer at lower utility values than the higher utility values. This pattern of distribution was consistent with other studies that revealed poor relationship between two instruments (26, 27). There are a number of reasons for these noticeable differences also, as stated by other authors (26, 27). Because of its high lower boundary, the SF-6D of 0.30 generates a narrower range of utility scores. This leads an underestimation in the changes of utility values for interventions that influence the lower end of the range. Another reason for the different results was the underlying techniques required to derive the utility algorithms. The EQ-5D utility algorithm is based upon the time trade-off (TTO) technique, whereas the SF-6D values were elicited using the standard gamble (SG) technique (5, 6). Some studies have demonstrated higher values for SG in comparison with TTO technique (28, 29).

This study estimated the utility values for different states of breast cancer that can be used in a model based cost-utility analysis and will be a valuable addition to the scientific literature. In addition, this is, to our best knowledge, the first study to report and compare the EQ-5D and SF-6D in an Iranian context in patients with breast cancer. Nevertheless, we compared the performance of two commonly used measures, and there appeared to be statistically and clinically significant differences in the utilities generated by these instruments. Consequently, if these values used as the weightings for QALYs, they would not result in comparable estimates. There are also some limitations to this study, which necessitate exercising caution when using results from this study. Firstly, the scoring algorithms for both instruments were based on a UK population, which may differ from values revealed by the Iranian population. Secondly, even though the patients were recruited consecutively, they were not assigned randomly. Third, the sample size used in the present study was quite small; therefore, further research with large samples will be required to confirm the findings from this study.

Acknowledgments

This study was part of a Ph.D. dissertation supported by the Iran University of Medical Sciences. The authors would like to thank the staff of the BCRC breast cancer subspecialty clinic in Tehran, Iran, for their great assistance in this project.

Footnotes

Authors' Contribution: Mahmood Yousefi participated in developing the study design, data acquisition, statistical analysis and drafting the manuscript. Safa Najafi provided technical support. Shahram Ghaffari revised the manuscript for important intellectual content. Hossein Ghaderi contributed to the statistical analysis, interpretation of the data, and drafting the manuscript.

Financial Disclosure: The authors declare that there are no conflicts of interest.

Funding/Support: This study was funded by the Iran University of Medical Sciences (IUMS) (Grant No: IUMS/2012/631).

References

1. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford University press; 2005.
2. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. *Applied methods of cost-effectiveness analysis in healthcare*. Oxford: Oxford University Press; 2010.
3. Brooks R. EuroQol: the current state of play. *Health Policy*. 1996;**37**(1):53-72. doi: 10.1016/0168-8510(96)00822-6. [PubMed: 10158943]
4. Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;**35**(11):1095-108. [PubMed: 9366889]
5. Dolan P, Roberts J. Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. *Med Care*. 2002;**40**(5):442-6. [PubMed: 11961478]
6. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;**21**(2):271-92. [PubMed: 11939242]
7. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol*. 1998;**51**(11):1115-28. [PubMed: 9817129]
8. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ*. 2004;**13**(9):873-84. doi: 10.1002/hec.866. [PubMed: 15362179]
9. Kopec JA, Willison KD. A comparative review of four preference-weighted measures of health-related quality of life. *J Clin Epidemiol*. 2003;**56**(4):317-25. [PubMed: 12767408]
10. Dolan P, Gudex C, Kind P, Williams A. *A social tariff for EuroQol*. UK: Centre for Health Economics University of York; 1995.
11. Brazier J, Deverill M, Green C. A review of the use of health status measures in economic evaluation. *J Health Serv Res Policy*. 1999;**4**(3):174-84. [PubMed: 10538884]
12. Ware JE, Kosinski M, Dewey JE, Gandek B. *SF-36 health survey: manual and interpretation guide*. Lincoln, US: Quality Metric Inc; 2000.
13. Longworth L, Bryan S. An empirical comparison of EQ-5D and SF-6D in liver transplant patients. *Health Econ*. 2003;**12**(12):1061-7. doi: 10.1002/hec.787. [PubMed: 14673814]
14. Joore M, Brunenberg D, Nelemans P, Wouters E, Kuijpers P, Honig A, et al. The impact of differences in EQ-5D and SF-6D utility scores on the acceptability of cost-utility ratios: results across five trial-based cost-utility studies. *Value Health*. 2010;**13**(2):222-9. doi: 10.1111/j.1524-4733.2009.00669.x. [PubMed: 19878492]
15. Marra CA, Esdaile JM, Guh D, Kopec JA, Brazier JE, Koehler BE, et al. A comparison of four indirect methods of assessing utility values in rheumatoid arthritis. *Med Care*. 2004;**42**(11):1125-31. [PubMed: 15586840]
16. Barton GR, Bankart J, Davis AC. A comparison of the quality of life of hearing-impaired people as estimated by three different utility measures. *Int J Audiol*. 2005;**44**(3):157-63. [PubMed: 15916116]
17. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;**127**(12):2893-917. doi: 10.1002/ijc.25516. [PubMed: 21351269]
18. Pickard AS, Wilke CT, Lin HW, Lloyd A. Health utilities using the EQ-5D in studies of cancer. *Pharmacoeconomics*. 2007;**25**(5):365-84. [PubMed: 17488136]
19. Matalqah LM, Radaideh KM, Yusoff ZM, Awaisu A. Health-related quality of life using EQ-5D among breast cancer survivors in comparison with age-matched peers from the general population in the State of Penang, Malaysia. *J Public Health*. 2011;**19**(5):475-80. doi: 10.1007/s10389-011-0406-6.
20. Montazeri A. Health-related quality of life in breast cancer patients: a bibliographic review of the literature from 1974 to 2007. *J Exp Clin Cancer Res*. 2008;**27**:32. doi: 10.1186/1756-9966-27-32. [PubMed: 18759983]
21. Rosner B. *Fundamentals of Biostatistics*. Belmont: Thomson Brooks Cole; 2006.
22. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998;**17**(1):101-10. [PubMed: 9463853]
23. Lidgren M, Wilking N, Jonsson B, Rehnberg C. Health related quality of life in different states of breast cancer. *Qual Life Res*. 2007;**16**(6):1073-81. doi: 10.1007/s1136-007-9202-8. [PubMed: 17468943]
24. Ware Jr JE. SF-36 health survey update. *Spine (Phila Pa 1976)*. 2000;**25**(24):3130-9. [PubMed: 11124729]
25. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res*. 2005;**14**(6):1523-32. [PubMed: 16110932]
26. van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary heart disease. *Health Qual Life Outcomes*. 2006;**4**:20. doi: 10.1186/1477-7525-4-20. [PubMed: 16563170]
27. Xie F, Li SC, Luo N, Lo NN, Yeo SJ, Yang KY, et al. Comparison of the EuroQol and short form 6D in Singapore multiethnic Asian knee osteoarthritis patients scheduled for total knee replacement. *Arthritis Rheum*. 2007;**57**(6):1043-9. doi: 10.1002/art.22883. [PubMed: 17665466]
28. Tsuchiya A, Brazier J, Roberts J. Comparison of valuation methods used to generate the EQ-5D and the SF-6D value sets. *J Health Econ*. 2006;**25**(2):334-46. doi: 10.1016/j.jhealeco.2005.09.003. [PubMed: 16271783]
29. Green C, Brazier J, Deverill M. Valuing health-related quality of life. A review of health state valuation techniques. *Pharmacoeconomics*. 2000;**17**(2):151-65. [PubMed: 10947338]