

Validity and reliability of the Arabic version of the the EuroQOL (EQ-5D). A study from Saudi Arabia

Abdulkareem M. Bekairy¹,
Rami T. Bustami¹,
Mohammed Almotairi¹,
Anan Jarab²,
Abdulmalik M. Katheri¹,
Tariq M. Aldebasi³,
Salaheddin Aburuz¹

¹Department of Pharmacy Practice, College of Pharmacy, King Saud Bin Abdul-Aziz University for Health Sciences, Riyadh, Saudi Arabia, ²Department of Clinical Pharmacy, College of Pharmacy, Jordan University of Science and Technology, Irbid, Jordan, ³Department of Surgery, King Abdulaziz Medical City, National Guard Health Affairs, Riyadh, Saudi Arabia

Address for correspondence:

Abdulkareem Bekairy,
College of Pharmacy, King Saud Bin Abdul-Aziz University for Health Sciences, Riyadh 11481, Kingdom of Saudi Arabia.
Tel.: +966 11 42 95017.
E-mail: Bekairy@ksau-hs.edu.sa

WEBSITE: ijhs.org.sa

ISSN: 1658-3639

PUBLISHER: Qassim University

ABSTRACT

Objectives: The EuroQOL five-dimension questionnaire (EQ-5D) is one of the most widely used tools to evaluate health-related quality of life (HRQOL). Studies on the psychometric features of the EQ-5D in the Gulf region of the Middle East are lacking. This study is aimed at evaluating the validity and reliability of the Arabic version of the EQ-5D in Saudi Arabia.

Methods: The study was done utilizing a convenience sample of Arabic-speaking adult patients or visitors at the outpatient clinic at King Abdulaziz Medical City (KAMC) in Riyadh during October–December 2015. Participants were interviewed using translated into Arabic versions of the EQ-5D and the Short Form Health Survey (SF-36). Known-group construct validity of the EQ-5D and visual analog scale (EQ-VAS) was assessed through testing a number of hypotheses comparing responses to EQ-5D dimensions or EQ-VAS to SF-36 scores by dimension and external variables. Test-retest reliability was also assessed.

Results: Eighty subjects were included in the study. Mean \pm standard deviation age was 33.0 ± 11.6 years, with 51% of females. Responses to EQ-5D dimensions revealed that the proportion of respondents reporting problems ranged from 6% for problems with self-care to 48% for pain/discomfort. All of the five hypotheses linking EQ-5D responses to external variables were satisfied. Test-retest reliability was evaluated using Cohen's κ , which ranged from 0.53 to 1.00.

Conclusion: Validity and reliability of the Arabic EQ-5D were established for evaluating HRQOL in Saudi Arabia. Further studies are warranted to evaluate the validity and reliability of this standard tool in other Arabic-speaking countries with varying cultures.

Keywords: Health-related quality of life, EQ-5D, SF-36, construct validity

Background

Health-related quality of life (HRQOL) is a comprehensive multidimensional concept that incorporates social, physical, and mental well-being.^[1] HRQOL has an important role in clinical settings. It has been utilized to evaluate humanistic and economic outcomes in different clinical trials for patients with various chronic diseases.^[2,3] Furthermore, HRQOL data have been shown to be useful to allocate resources.^[4]

Numerous tools have been used to evaluate HRQOL. The EuroQOL five-dimension questionnaire (EQ-5D) is one of the most commonly used questionnaires to assess HRQOL. The EQ-5D is a standardized tool for use as a measure of health outcome. The EQ-5D assesses a broad notion of health including the medical classification and the physical, emotional, and social functioning. It incorporates both favorable

(wellbeing) and unfavorable aspects (sickness) of health. In addition to the fact that the EQ-5D is user-friendly, short, and flexible, it comprises a questionnaire and a visual analog scale (EQ-VAS), a self-rated health tool that can be utilized to detect variations over time. The questionnaire allows the subject to provide a detailed description of their present health in 5 dimensions such as self-care, mobility, usual activities, pain/discomfort, and anxiety/depression. The EuroQOL foundation has lately developed several new surveys to further improve the EQ-5D questionnaire; some of which were utilized to assess a version of the EQ-5D that includes 5 levels (EQ-5D [5L]) and the newly-developed EQ-5D-Y version for children.

The English version of the EQ-5D was used in several studies that assessed HRQOL.^[5-7] It has been extensively validated and proved to be internally consistent and reliable for use in the general population as well as in specific

patient groups.^[5,8] In 1993, the EQ-5D was available in only 6 languages. It has now been translated into 63 languages, and new versions currently being developed including versions for Indonesia, Malaysia, Taiwan, Thailand, South Africa, Germany, and the Philippines. The translated versions have been shown to be valid and reliable in the patient and general population.^[9-13]

There is only a few studies that were focused on the assessment of the validity and reliability of the Arabic EQ-5D. The Arabic language is among the top ten languages in terms of the number of native speakers worldwide ranking fifth with an estimated 300 million speakers representing 4.5% of the world population. A recent census estimated that the number of Arabic-speaking people exceeded 2 million in the US alone (Arab American Institute Foundation, 2012).^[14] An earlier study was conducted in Jordan to assess the validity and reliability of the Arabic EQ-5D.^[4] This was the first study conducted in an Arabic-speaking country using the Arabic EQ-5D. The study demonstrated that the Arabic EQ-5D is valid and reliable in assessing HRQOL among patients or visitors at an outpatient clinic in Jordan.

To the best of our knowledge, studies on the psychometric features of the EQ-5D in the Gulf region are lacking due to language and cultural barriers. Saudi Arabia is by far the largest Arabic-speaking country in the Gulf region and is one of the largest Arabic-speaking countries in the Middle East, with an estimated population of more than 30 million.^[15] The aim of this study is to evaluate the validity and reliability of the Arabic EQ-5D in a regional academic medical center in Riyadh, Saudi Arabia.

Methods

The study was done using a sample of Arabic-speaking adult patients or visitors at the outpatient clinic at King Abdulaziz Medical City (KAMC) in Riyadh during October–December 2015. The study was approved by the Ethics Committee at KAMC. The Arabic EQ-5D and the Arabic short form health survey (SF-36) were administered to all participants who were interviewed by trained research assistants. Demographic data were also obtained using a standard questionnaire. Test-retest reliability was assessed by administering the Arabic EQ-5D to a random sample of the participants twice, 3 weeks apart.

The five dimensions of the EQ-5D are mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.^[12] The SF-36 comprises 36 questions that assess HRQOL on 8 dimensions including physical functioning, mental health status, role-physical, general health status, bodily pain, vitality, social functioning, and role emotional.^[16] The EQ-VAS is a scale calibrated from 0 to 100 with higher scores indicating better health status.^[12]

Known-group construct validity of the EQ-5D and EQ-VAS was assessed through testing several hypotheses that compare

responses to EQ-5D dimensions or EQ-VAS to SF-36 scores. The following five hypotheses were used to test the validity and reliability of the Arabic EQ-5D:

1. Subjects who have a problem in any EQ-5D dimension have decreased scores on all SF-36 dimensions.^[4,17]
2. Subjects who have some problem in EQ-5D physical dimension have lower SF-36 scores in the physical functioning dimension.^[4]
3. Subjects who have some problem in EQ-5D anxiety/depression dimension have lower SF-36 scores in the mental health dimension.^[4]
4. Subjects with poor general health would have lower scores in EQ-5D indexes and VAS.^[4]
5. Females have lower EQ-VAS scores compared to males.^[4]

Statistical analysis

Descriptive statistical analyses were performed for the study participants. Continuous variables were summarized using mean, standard deviation (SD), median, and interquartile range (IQR). Proportions were used for categorical variables. Hypothesized trends were tested using the chi-square test or Fisher's exact test for categorical variables and *t*-test or Mann–Whitney *U*-test for continuous variables. Cronbach's alpha (α) was used to quantify internal consistency of the EQ-5D, and test-retest reliability was examined using Cohen's Kappa (κ). Statistical significance was considered at $P < 0.05$. All statistical analyses were performed using SPSS 21.0 (Release 21.0.0.0, SPSS Inc., USA).

Results

A total of 80 subjects were included in the study. Descriptive statistics of the study participants are displayed in Table 1. Mean \pm SD age was 33.0 ± 11.6 years, with 51% of females. The majority of respondents had an educational level of high school or above (76%). Mean \pm SD EQ-VAS was 74.9 ± 17.6 .

The distribution of responses to EQ-5D dimensions is shown in Table 2. The results showed that the proportion of respondents

Table 1: Profile of respondents $n=80$

Factor	Value
Gender n (%)	
Female	41 (51)
Male	39 (49)
Age (years) mean \pm SD	33.0 \pm 11.6
Median (IQR)	31.5 (23.3–40)
Educational level n (%)	
Elementary/secondary	19 (24)
High school or above	61 (76)
EQ-VAS score mean \pm SD	74.9 \pm 17.6
Median (IQR)	80 (70–85)

SD: Standard deviation

reporting problems was as follows: 48% reported problems with pain/discomfort, 44% with mobility, 44% with anxiety/depression, 35% with usual activities, and only 6% reported problems with self-care.

The Cronbach's α value was 0.72, suggesting that the Arabic EQ-5D has an acceptable level of internal consistency. The Arabic EQ-5D was administered twice to 24 respondents (30%), with a median follow-up time of 3 weeks (interquartile range: 2–4 weeks). The value of Cohen's κ for the four EQ-5D dimensions was 0.80, 1.0, 0.81, 0.53, and 0.84 for mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, respectively ($P < 0.001$ for all dimensions).

Results in Table 3 showed that respondents who reported a problem in any of the EQ-5D dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) had significantly lower scores for most SF-36 scales ($P < 0.05$). Furthermore, significantly lower scores on the SF-36 physical functioning and mental health scales were observed for

respondents reporting problems in the EQ-5D physical and anxiety/depression dimensions ($P = 0.001$ and $P < 0.001$, respectively). Subjects with poor general health as indicated by the SF-36 had significantly lower EQ-VAS scores and scores for most EQ-5D dimensions ($P < 0.05$) [Figure 1]. Finally, females had significantly reduced EQ-VAS scores compared to males ($P = 0.021$) [Figure 2].

DISCUSSION

Several studies have focused on testing the validity and reliability of the EQ-5D in regions like Europe where various languages are spoken.^[11,17] This is the first study to establish the validity and reliability of the EQ-5D instrument in the Saudi population. All of the hypotheses linking the Arabic EQ-5D to SF-36 were satisfied, suggesting that the Arabic EQ-5D has characteristics that are comparable to those of

Table 2: Responses to EQ-5D dimensions. Number of respondents=80

EQ-5D dimension	Response n (%)		
	No problem	Moderate problem	Extreme problem
Mobility	45 (56.3)	35 (43.8)	0 (0)
Self-care	75 (93.8)	5 (6.3)	0 (0)
Usual activities	52 (65.0)	26 (32.5)	2 (2.5)
Pain/discomfort	42 (52.5)	36 (45.0)	2 (2.5)
Anxiety/depression	45 (56.3)	33 (41.3)	2 (2.5)

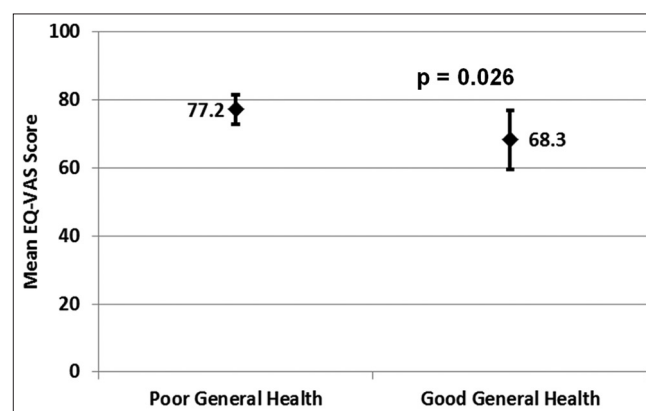


Figure 1: Mean (95% CI) EQ-VAS scores by general health status. Number of respondents = 80

Table 3: Mean (SD) SF-36 scores for respondents in the five dimensions of the EQ-5D. Number of respondents=80

EQ-5D Dimension	N	PF	RP	BP	GH	VT	SF	RE	MH
Mobility									
No problem	45	85.9 (16.9) ^a	75.0 (36.1) ^a	74.3 (32.9) ^b	59.7 (19.4) ^b	58.3 (17.7) ^b	75.3 (23.2)	63.0 (42.8) ^b	67.6 (18.0)
With problem	35	64.7 (20.0)	37.9 (42.2)	55.7 (31.4)	50.5 (17.5)	49.3 (18.6)	65.7 (20.4)	41.0 (43.6)	62.9 (19.2)
Self-care									
No problem	75	78.5 (19.6) ^b	59.7 (42.5)	67.7 (32.9)	56.8 (18.6) ^b	55.8 (17.5) ^b	71.5 (22.2)	55.6 (43.6)	66.8 (17.6) ^b
With problem	5	49.0 (25.6)	45.0 (51.2)	43.0 (35.3)	38.4 (19.5)	33.0 (22.8)	65.0 (27.1)	20.0 (44.7)	45.6 (23.3)
Usual activities									
No problem	52	84.3 (16.6) ^a	72.6 (37.8) ^a	74.7 (30.7) ^a	60.6 (18.1) ^a	60.5 (15.1) ^a	75.0 (22.7) ^b	61.5 (43.5) ^b	69.2 (15.7) ^b
With problem	28	62.3 (21.2)	33.0 (40.3)	50.4 (32.9)	46.5 (17.5)	43.0 (19.2)	63.8 (20.2)	38.1 (42.3)	58.7 (21.6)
Pain/discomfort									
No problem	42	83.7 (19.8) ^a	80.4 (34.7) ^a	86.7 (23.8) ^a	60.9 (18.2) ^a	63.9 (14.0) ^a	79.5 (20.3) ^a	62.7 (43.1) ^b	72.2 (15.0) ^b
With problem	38	68.8 (19.8)	34.9 (38.4)	43.6 (27.4)	49.9 (18.4)	43.8 (17.3)	61.8 (21.1)	43.0 (43.8)	58.1 (19.4)
Anxiety/depress.									
No problem	45	78.1 (20.3)	68.3 (40.0) ^b	76.2 (29.9) ^b	61.9 (16.9) ^b	61.0 (15.0) ^a	77.5 (22.4) ^a	70.4 (39.7) ^a	74.4 (12.4) ^a
With problem	35	74.7 (22.2)	46.4 (43.8)	53.3 (33.5)	47.7 (18.8)	45.9 (19.3)	62.9 (19.8)	31.4 (40.4)	54.1 (19.0)

^a $P < 0.05$, ^b $P < 0.001$: PF: Physical functioning, RP: Role limitation due to physical problem, BP: Bodily pain, GH: General health, VT: Vitality, SF: Social functioning, RE: Role limitation due to emotional problem, MH: Mental health

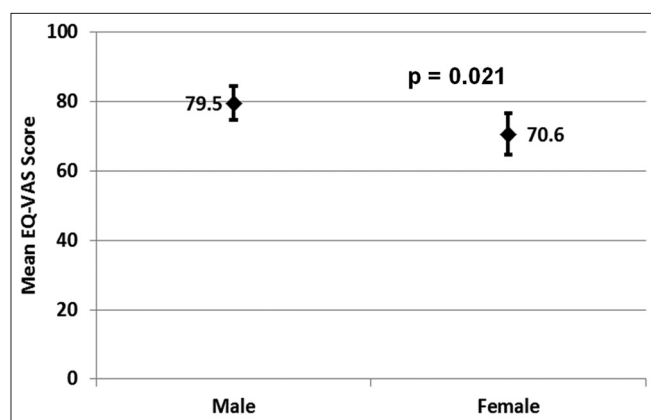


Figure 2: Mean (95% CI) EQ-VAS scores by gender. Number of respondents = 80

other EQ-5D versions. These results support the findings from a relatively recent study conducted in Jordan using the same instrument.^[4]

We also found strong supporting evidence of the reliability of the Arabic EQ-5D self-classifier, with a Cohen's κ range of 0.53–1.00 indicating moderate-to-excellent agreement for the mobility, self-care, usual activities, anxiety/depression, and pain/discomfort dimensions. The Cohen's κ values in our study generally indicate that the Arabic EQ-5D is reliable in measuring HRQOL in Saudi Arabia.

Our results have demonstrated that the most common problems reported by respondents were pain/discomfort and anxiety/depression. The latter is usually associated with various health problems including diabetes, cardiovascular diseases, and hypertension. This finding is significant as these are common health problems among the Saudi population where the estimated prevalence of diabetes and hypertension is 30% and 26%, respectively.^[18,19]

A strength of this study is its use of widely used instruments to measure HRQOL among Arabic-speaking patients. Another important outcome of this study was its value in terms of exploring and validating the Arabic EQ-5D through testing several hypotheses that compare responses to EQ-5D dimensions or EQ-VAS to SF-36 scores. This finding is extremely valuable and set the stage for utilizing this important tool for measuring HRQOL among patients in numerous countries in the Gulf region and the Middle East where the Arabic language is widely spoken.

Our study has a number of limitations. First, we used a convenience sample of a relatively small size, which may restrain the generalizability of our findings. This limitation was apparent as our sample of respondents was relatively young with a high level of education. Nevertheless, as stated above, our results are consistent with those reported in a similar study that utilized the Arabic EQ-5D in Jordan.

Furthermore, this is a pilot study, and hopefully a future investigation will be conducted including a larger and more representative sample from both the inpatient and outpatient setting with more variant demographic and clinical characteristics.

Conclusion

The Arabic EQ-5D was shown to be valid and reliable in evaluating HRQOL in Saudi Arabia and thus can be used as a tool for patients whose primary language is Arabic. Further studies are warranted to assess the validity and reliability of this standard tool in other Arabic-speaking countries with varying cultures.

References

1. Available from: <http://www.cdc.gov/hrqol>. [Last accessed on 2016 Apr 18].
2. Higginson IJ, Carr AJ. Measuring quality of life: Using quality of life measures in the clinical setting. *BMJ* 2001;322:1297-300.
3. Hennessy CH, Moriarty DG, Zack MM, Scherr PA, Brackbill R. Measuring health-related quality of life for public health surveillance. *Public Health Rep* 1994;109:665-72.
4. Aburuz S, Bulatova N, Twalbeh M, Gazawi M. The validity and reliability of the Arabic version of the EQ-5D: A study from Jordan. *Ann Saudi Med* 2009;29:304-8.
5. Isitt JJ, Nadipelli VR, Kouassi A, Fava M, Heidbreder C. Health-related quality of life in acute schizophrenia patients treated with RBP-7000 once monthly risperidone: An 8-week, randomized, double-blind, placebo-controlled, multicenter phase 3 study. *Schizophr Res* 2016;174:126-31.
6. Luo N, Chew LH, Fong KY, Koh DR, Ng SC, Yoon KH, *et al.* Validity and reliability of the EQ-5D self-report questionnaire in english-speaking Asian patients with rheumatic diseases in Singapore. *Qual Life Res* 2003;12:87-92.
7. Luo N, Low S, Lau PN, Au WL, Tan LC. Is EQ-5D a valid quality of life instrument in patients with parkinson's disease? A study in singapore. *Ann Acad Med Singapore* 2009;38:521-8.
8. Korman NJ, Zhao Y, Pike J, Roberts J. Relationship between psoriasis severity, clinical symptoms, quality of life and work productivity among patients in the USA. *Clin Exp Dermatol* 2016;41:514-21.
9. Varatharajan S, Chen W-S. Reliability and validity of EQ-5D in Malaysian population. *Appl Res Qual Life* 2012;7:209-21.
10. Chang TJ, Tarn YH, Hsieh CL, Liou WS, Shaw JW, Chiou XG, *et al.* Taiwanese version of the EQ-5D: Validation in a representative sample of the taiwanese population. *J Formos Med Assoc* 2007;106:1023-31.
11. Stark RG, Reitmeir P, Leidl R, Konig HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. *Inflam bowel dis* 2010;16:42-51.
12. Luo N, Chew LH, Fong KY, Koh DR, Ng SC, Yoon KH, *et al.* Validity and reliability of the EQ-5D self-report questionnaire in chinese-speaking patients with rheumatic diseases in singapore. *Ann Acad Med Singapore* 2003;32:685-90.
13. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burström K, Cavrini G, *et al.* Feasibility, reliability, and validity of the EQ-5D-Y: Results from a multinational study. *Qual Life Res* 2010;19:887-97.
14. Available from: <http://www.aaiusa.org>. [Last accessed on 2016 May 22].

15. Available from: <http://www.countrymeters.info/en/Saudi-Arabia>. [Last accessed on 2016 May 22].
16. Wagner AK, Gandek B, Aaronson NK, Acquadro C, Alonso J, Apolone G, *et al.* Cross-cultural comparisons of the content of SF-36 translations across 10 countries: Results from the international quality of life assessment project. *J Clin Epidemiol* 1998;51:925-32.
17. Brazier J, Jones N, Kind P. Testing the validity of the euroqol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res* 1993;2:169-80.
18. Alqurashi KA, Aljabri KS, Bokhari SA. Prevalence of diabetes mellitus in a Saudi community. *Ann Saudi Med* 2011;31:19-23.
19. Al-Nozha MM, Abdullah M, Arafah MR, Khalil MZ, Khan NB, Al-Mazrou YY, *et al.* Hypertension in Saudi Arabia. *Saudi Med J* 2007;28:77-84.