

# Cross-Cultural Adaptations and Psychometric Properties of the Quality of Life Scales used for the Adult Arabic-Speaking Population: A Systematic Review

## Abstract

The aim of this review is to explore the psychometric properties of Arabic Quality of Life scales to identify the appropriate scales for use in research and clinical practice. A systematic search was conducted using four databases; CINAHL, MEDLINE, EMBASE, and PsycINFO. This review was carried out following the Preferred Reporting Items Systematic Reviews and Meta-Analysis (PRISMA) guidelines, and the quality assessment procedure used by Terwee et al. was utilised to evaluate the psychometric properties. There were 27 studies found relating to the psychometric properties of seven different scales. The studies provided sufficient information about validity and reliability, but not all studies reported translation and cross-cultural adaptation processes. Seven scales were identified for use in the investigation of the QOL among the Arabic speaking population. It is suggested that researchers and clinicians consider which scales can facilitate and measure the subscales and sample characteristics required for their population of interest.

**Keywords;** Quality of Life; Medicine, Arabic; Cross-Cultural Comparison; Psychometrics; Systematic Review; Surveys and Questionnaires

## **Introduction:**

The concept of Quality of Life (QoL) is important because it measures the effect of diseases in patients in research and in clinical practice. In addition, QoL is an essential endpoint in the treatment plan for the policy maker, the healthcare provider and the patient. The impact of medical intervention on functioning, lifestyle and well-being have increased interest in the scales that measure QoL.<sup>1</sup> It is considered to be a multidimensional construct which generally relies on personal characteristics and contextual and environmental variables.<sup>2</sup> As a consequence, it is necessary to identify robust scales with good psychometric properties and cross-cultural adaptations that can be used to measure QoL.

Most QoL assessment scales are available in English, and the demand to translate and adapt them for different languages and cultures has increased. Cross-cultural adaptation and translation is a systematic process that prepares the scales for use in another setting.<sup>3</sup> However, the scale should maintain its content validity after translation and cultural adaptation.

There are approximately 420 million Arabic-speaking people living in 23 countries.<sup>4</sup> There are two forms of the Arabic language. The first is modern standard Arabic which is used in the written form, in official settings and in education. The second form is the regional dialect which is colloquial.<sup>5</sup> The Arabic version of QoL measures were not available until Coons *et al.*<sup>6</sup> conducted the first psychometric study in order to translate and validate the Arabic version of the QoL scales SF-36.

Since then, many different QoL scales have been translated. Subsequently, there is a need to evaluate these and identify which have good cross-cultural adaptation and good quality validation. The aim of this review is to explore the psychometric properties of Arabic QoL scales, in order to identify appropriate scales that can be used in research and clinical practice.

## **Methods:**

This systematic review has been carried out to identify the robust QoL scales that can be used with Arabic populations. The review has been conducted according to the PRISMA guidelines.<sup>7</sup>

## **Search Strategy:**

Studies investigating the QoL involving Arabic participants were identified by searching the following electronic databases; CINAHL, MEDLINE, EMBASE and PsycINFO, for the period

from 1946 to April 2019. Search terms used **combination** of free text words and MeSH terms in each database. These included; "Psychometrics" OR "Reliability" OR "Validity" OR "Instrument Validation" **AND** "Arabs" OR "Medicine, Arabic" **AND** "Functional Status" OR "well-being or wellbeing or well being " OR "Quality of Life " OR "Health Status " OR "Health and Life Quality " OR "Quality of Health Care " OR "Assessment" OR "Patient Assessment " OR "Clinical Assessment Tools" OR "Health Impact Assessment" OR "Clinical Assessment Tools" OR "Outcome Assessment" OR "measurement tool" OR "Questionnaires". In addition, reference lists were screened to identify any further studies.

### **Study Selection:**

The first author (AM) conducted the search and assessed all the papers to determine their eligibility. Studies were considered eligible for inclusion; (1) if they were published in English, 2) involved adult patients above the age of 18 years, 3) were psychometric studies only with information about validity or reliability, 4) involved QoL measures which were translated into Arabic, 5) involved the Arabic-speaking population, and 6) had no restrictions on the study design. The exclusion criteria were studies that measured the QoL for one specific disease only, and were not psychometric studies.

### **Quality Assessment:**

#### **Quality Criteria**

Using the quality assessment criteria of Terwee *et al.*, the psychometric properties of the scales were evaluated by the lead author (AM).<sup>8</sup> These include content validity, internal consistency, criterion validity, construct validity, reproducibility, responsiveness, floor and ceiling effects, and interpretation (Table 1). The nine aspects were given various ratings, including; '+' positive rating; '?' indeterminate rating; '-' negative rating; and, '0' no information available. Terwee *et al.* recommend presenting the results in a table but not using an overall quality score to provide an overview of all the scales.<sup>8</sup> The overall score appears to indicate equal importance for each psychometric property, but in practice this is not the case.

Reliability refers to the reproducibility or consistency of scores from one assessment to another.<sup>9</sup> Internal consistency is usually reported as a coefficient ranging from 0 to 1, (a value of 0 represents no correlation, whereas a value of 1 represents a perfect correlation). An acceptable value is 0.7, while more than 0.9 is considered highly reliable.<sup>10</sup> Other reliability

measures are the inter-rater, intra-rater or test-retest. Validity is the ability of the scale to measure the attributes of the construct under consideration, in other words the degree to which the scale measures what it is intended to measure. Validity is divided into three types; content, construct, and criterion (concurrent and predictive).<sup>11</sup>

### **Cross-Cultural Adaptation and Translation Criteria**

Cross-cultural adaptation and translation of the scales were evaluated based on the guidelines suggested by Guillemin *et al.*<sup>12</sup> The guidelines suggest five steps to the sense of the scale in terms of the target culture; (1) translation, (2) back-translation, (3) committee review, (4) pre-testing, and (5) re-examination of the weighing of scores.

With regard to the translation, it is recommended that at least two qualified translators should be used to translate the scale from the original language to the target language. Back-translations use two independent translators who translate the version in the target language back to the language of the original version. This step ensures that the translated version reflects the content of the original scale. The third step involves a committee review to develop the pre-final version for pre-testing. The fourth step is the pre-testing which tests the pre-final version on 30-40 subjects or patients from the target population. The final step is the weighting of the scores, which considers the weights of the scores in the cultural context. Each step is assessed using the following scoring categories; 1 = poor, 2 = moderate and 3 = good.<sup>12</sup> The overall score is the mean of the values obtained for each of the five steps.

### **Results:**

The database search identified 1082 articles (Figure 1). After screening the titles and abstracts against the inclusion and exclusion criteria, the number of articles was reduced to 38. After reading the full text, 27 papers met all the inclusion criteria. All 27 studies were published between 1998 and 2019.

Seven self-reporting scales were identified that were translated and tested psychometrically in Arabic. The Medical Outcomes Study Short Forms (SF-36 versions, SF-12) were investigated in ten studies, the Dartmouth CO-OP Functional Health Assessment Charts/World Organisation of Family Doctors (CO-OP/WONCA charts) in one study, the World Health Organisation Quality of Life Brief version (WHOQOL-brief) in three studies, the EuroQol Group health status index 5-Dimensions (EQ- 5D) in one study, the European Organisation for Research and

Treatment of Cancer (EORTC) (all specific with general quality of life) (EORTC QLQ-C30, EORTC QLQ-C15-PAL) in five studies, the Functional Assessment of Cancer Therapy-General FACT- (G) in two studies, and the Quality of Life Index (QLI) in one study. The EORTC QLQ- C30 and the FACT-G were included in this review as they are used for multiple types of cancer.

Seven scales were excluded from the review because they were developed and validated for a specific disease, so cannot be used in other populations. For example, allergic-rhinitis <sup>13</sup>, Behçet's disease <sup>14</sup>, Alzheimer's disease <sup>15</sup>, schizophrenia <sup>16</sup>, ENT head and neck cancer <sup>17,18</sup>, and breast cancer <sup>19</sup>.

### **Study Characteristics:**

The studies considered in this review were conducted in different Arab-speaking countries. Seven studies were undertaken in Jordan, four studies in Saudi Arabia, two each in Egypt, Morocco, Kuwait, Tunisia, Lebanon and the United Arab Emirates (UAE) and one each in Sudan and Qatar. However, two of the studies were conducted in the Netherlands which is not an Arabic-speaking country, but the sample used was Moroccan Arabic (table 2).

Furthermore, 21 studies were translated into modern standard Arabic which can be used in all Arabic populations. Six studies were translated into Arabic language dialect which can only be used with specific Arabic populations. Three studies were translated into Moroccan Arabic dialect,<sup>20-22</sup> two into Tunisian Arabic dialect,<sup>23,24</sup> and one into Egyptian Arabic dialect.<sup>25</sup>

The studies in this review used quantitative research methods; 20 were cross-sectional surveys and seven were longitudinal. Cross-sectional data was collected at a specific point in time without follow-up, while longitudinal data was collected over different periods of time <sup>26</sup>. The longitudinal studies give more precise details about the temporal changes or treatment effects that can have an impact on QoL. Cross-sectional designs have difficulty in terms of creating a cohesive narrative on the impact of treatment on the QoL of patients. No studies used a mixed method approach.

### **Assessment of Psychometric Properties:**

The psychometric properties of the scales are presented in Table 3. None of the studies in this review tested all nine psychometric aspects as suggested by Terwee *et al.* <sup>8</sup>

Twenty-two studies achieved a positive score in terms of content validity. Five studies obtained a negative rating as the paper did not give any information about content validity. The internal consistency was tested in most of the studies, and in 26 studies there was a positive score. Only one study did not report the internal consistency.<sup>21</sup> The validity criterion was tested in two studies and achieved a positive rating,<sup>27,28</sup> while 25 studies did not provide any information with regard to criterion validity.

Construct validity was assessed in 22 studies; 21 were scored as positive, and only one study<sup>24</sup> scored as intermediate. Information about reproducibility-agreement was present in one study with a positive score.<sup>23</sup> Reliability was investigated in nine studies, seven of which had a positive score and two an intermediate score.<sup>29,30</sup> Two studies provided information about responsiveness, and in the study by Hoopman *et al.*<sup>22</sup> the score was positive, whereas it was intermediate in Hoopman *et al.*<sup>21</sup>.

Floor and ceiling effects were tested in six studies with four scoring positive<sup>20,22,27,31</sup> and two scoring intermediate.<sup>21,30</sup> Only Aburuz *et al.*<sup>13</sup> reported interpretation with an intermediate score, with 26 studies not providing any information.

The SF-36 demonstrated the most robust psychometric quality rating, followed by the WHOQOL-Brief. Only the SF-36 tested seven items recommended by the quality assessing criteria of Terwee *et al.* and reported positive ratings in terms of content validity, internal consistency, criterion validity, construct validity, agreement, reliability and floor and ceiling effects.<sup>8</sup> The content validity, internal consistency, and construct validity was rated as positive in the WHOQOL-Brief, but reliability and floor and ceiling effects were rated as indeterminate.

### **Translation Assessment:**

The process of cross-cultural adaptation and the translation of the scales into an Arabic version was not reported in every study. This may be because the scales had been translated into Arabic in earlier studies. The evaluation of the Guillemin *et al.* guidelines in each study is presented in Table 4.<sup>12</sup> The review found that fourteen studies used translation and cross-cultural adaptation of the QoL scales.

Only two studies, Coons *et al.*<sup>6</sup> and Soudy *et al.*<sup>32</sup> adopted all five steps of the guidelines, while ten studies reported four of the steps without providing any information about the weighting score adaptation step. Al Barmawi *et al.*<sup>33</sup> presented a three step translation technique and a

back-translation technique and pre-testing, while Hoopman *et al.*<sup>22</sup> reported the first two steps. Only Alawneh *et al.*<sup>34</sup> reported a one-step translation technique.

The cross-cultural adaptation of the EORTC-QLQ, FACT-(G) and QLI scales all met the quality criteria and received an overall mean score of three. One scale scored 2.5 (CO-OP/WONCA). However, there was no information about the cross-cultural adaptation given in the studies that validated the WHOQOL-Brief scales.

## **Quality of Life Scales:**

### **EORTC-QLQ-C30**

The European Organisation for Research and Treatment of Cancer, Quality of Life Questionnaire (EORTC-QLQ) has developed several scales to assess the QoL of cancer patients. The EORTC-QLQ-C30 consists of nine multi-item scales; five functional sub-scales (physical, role, cognitive, emotional, and social), three symptom sub-scales (fatigue, pain, and nausea and vomiting), and a global health status and quality of life scale.<sup>35</sup> An additional six single items assessing the symptoms of cancer patients were included (dyspnoea, insomnia, appetite loss, constipation, diarrhoea and financial difficulties). The first 28 items of the scale used a 4-point Likert scale on a response scale of 'not at all' (1), to 'very much' (4). The 29 and 30 items were assessed using a 7-point numeric rating scale.

Four studies validated the EORTC-QLQ-C30 and were translated into modern standard Arabic.<sup>34,36-38</sup> All studies were cross-sectional in nature and included a total of 1313 cancer patients. The internal consistency as measured by the coefficient alpha ranged from < 0.70 to 0.91. The scale showed satisfactory psychometric properties. The psychometric testing occurred with regard to Arab cancer populations, consistent with the purpose of the scale. Aaronson *et al.*<sup>35</sup> published the first version of EORTC-QLQ-C30 which scored  $\geq 0.70$ , and similar results were reported in the Arabic version. The scale does not include spiritual and existential aspects.

The 30 items of EORTC-QLQ were reduced to 15 for Palliative Care to develop the EORTC-QLQ-C15-PAL.<sup>39</sup> The EORTC-QLQ-C15-PAL includes three multi-item scales; functional sub-scales (physical and emotional), symptom sub-scales (fatigue and pain), and a global health status and quality of life scale. The first 14 items of the scale used a 4-point Likert scale on a response scale of 'not at all' (1), to 'very much' (4). Item 15 was assessed using a 7-point numeric rating scale. Alawneh *et al.* translated and investigated the validity and reliability of

the EORTC-QLQ-C15-PAL using 175 Jordanian mixed cancer patients.<sup>34</sup> The internal consistency of the study was between the coefficient alpha < 0.71 and 0.90.

### **FACT-G**

The Functional Assessment of Cancer Therapy-General (FACT-G) is a 27-item, 5-point Likert scale which was originally validated using mixed cancer patients.<sup>40</sup> The scale assesses four dimensions; physical well-being, social/family well-being, and functional well-being with seven items, and emotional well-being with six items. In addition to cancer, the scale was used and validated on other chronic illness conditions in the general population. The scale also has specific items that can be used with the addition of general scales to measure specific types of cancer patients.

Four psychometric studies were conducted to evaluate the psychometric properties of the FACT-G in Arabic, with additional sub-scales. Lazenby *et al.* validated the FACT-G and a spiritual sub-scale using 205 mixed cancer patients.<sup>41</sup> Soudy *et al.*<sup>32</sup> validated the FACT-G and bone marrow transplantation sub-scales in 108 lymphoma patients who underwent stem cell transplantation, while Zahran *et al.*<sup>42</sup> validated the FACT-G and bladder cancer sub-scale in 90 bladder cancers; both studies were translated into modern standard Arabic. One study by Al Barmawi *et al.* assessed the psychometric properties of FACT-G in head-and/or-neck cancer patients.<sup>33</sup> All four studies were translated into modern standard Arabic. The internal consistency of FACT-G (Cronbach's  $\alpha$  range, 0.76 - 0.94), demonstrates almost the same internal consistency results as the original study by Cella *et al.*<sup>40</sup> The test-retest reliability has not yet been tested in the Arabic version. Further use of the scales in different clinical samples, and their psychometrics properties, need to be evaluated.

### **SF-36**

The Short Form 36 (SF-36) is a 36-item multi-purpose health survey. The scale consists of eight sub-scales including physical functioning, emotional problems, physical problems, mental health, bodily pain, general health, social functioning, vitality, and one single-item scale on health transition. The scale score ranged from 0 to 100, with a high score indicating a better quality of life and a low score a lower quality of life. The reliability of the original SF-36 exceeded 0.8.<sup>43</sup>



Nine studies tested the psychometric properties of the SF-36, and one study evaluated the shorter form of the SF-12.<sup>24</sup> First, the Arabic version of the SF-36 was validated and culturally adapted for Saudi Arabic scientific use by Coons *et al.*<sup>6</sup> The SF-36 was tested on multiple populations; four studies involved the general population<sup>6,20,23,27,31</sup>; one study involved burn patients<sup>25</sup>; one involved cancer patients<sup>22</sup>, ICU patients were included in another<sup>44</sup> and one was on Khat chewers<sup>45</sup>. The total sample size in all the studies was 2521. The scale was translated into three different Arab dialect languages. There were two studies in the Arab Moroccan dialect<sup>20,22</sup> and one study in both the Arab Egyptian dialect<sup>25</sup> and the Tunisian dialect<sup>23</sup>. An additional five studies were translated into standard Arabic. The internal consistency of all the studies ranged from 0.78 to 0.92. The test-retest reliability was assessed in four studies, with the intraclass correlation coefficient (ICC) exceeding 0.70. Younsi and Chakroun translated and investigated the validity and reliability of the Tunisian version of SF-12 in a Tunisian sample of 3582 individuals.<sup>24</sup> The coefficient alpha consistency was 0.73.

### **EQ-5D**

The EuroQol group design incorporated a 5-level scale to measure the QoL (EQ-5D).<sup>46</sup> The scale consists of 5-items that reflect five dimensions of QoL (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Each dimension has three responses: no problems, some/moderate problems, and extreme problems. In addition, health states were measured using a visual analogue scale ranging from 0 to 100.

Two studies tested the psychometric properties of the EQ-5D in Arabic. Aburuz *et al.*<sup>13</sup> translated and investigated the validity and reliability of the EQ-5D using a sample of 186 Jordanian patients from the general population, while Bekairy *et al.*<sup>47</sup> included 80 mixed Arabic patients. The scale was translated to modern standard Arabic which can be used with any Arabic-speaking people. The scale appears to be valid and reliable with an internal consistency of coefficient alpha  $\geq 0.72$ . The test-retest reliability of the scale using Cohen's  $\kappa$  for Aburuz *et al.*<sup>13</sup> ranged from 0.48 to 1.0, and for Bekairy *et al.*<sup>47</sup> between 0.53 and 1.00.

### **WHOQOL-Brief**

The World Health Organisation Quality of Life-Brief (WHOQOL-Brief) is a 26-item, 5-point Likert scale which was originally validated to measure people with a disease in the general population.<sup>48</sup> The WHOQOL-Brief has four sub-scales; physical health, psychological health,

social relationships, environmental health and two overall QoL and general health items. The original internal consistency ranged from 0.66 to 0.84.

Four studies were conducted to validate the WHOQOL-Brief in different countries; Ohaeri *et al.*<sup>29</sup> in Sudan; Ohaeri and Awadalla<sup>30</sup> in Kuwait; Bani-Issa<sup>49</sup> in the UAE; and Dalky *et al.*<sup>28</sup> in Jordan. The total sample size of all studies was 4392 and included psychiatric and diabetic patients and the general population. The internal consistency of the sub-scales ranged from 0.69 to 0.93, which indicated that the scales have an acceptable level of construct validity and reliability in terms of internal consistency. The Ohaeri and Awadalla test-retest reliability statistic (0.95) was significant.<sup>30</sup>

### **CO-OP/WONCA charts**

The Dartmouth Primary Care Cooperative Research Network (CO-OP) and the World Organisation of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians (WONCA) developed a QoL scale with six core functional status items including physical fitness, feelings, daily activities, social activities, change in health and overall health.<sup>50</sup> These items are assessed using a 5-point Likert scale.

The scale was translated and culturally adapted by Hoopman *et al.* to Moroccan Arabic (Tarifit) which is a local dialect in Morocco.<sup>21</sup> The sample was 37 mixed cancer patients. The scale was found to have adequate feasibility and construct validity, but discriminant validity could only be partially confirmed. The sample size was small and used a local language group only, so further examination of the psychometric properties is required for modern Arabic.

### **QLI**

The Quality of Life Index (QLI) consists of 70 items that assess health and functioning, socio-economic, psychological/spiritual, and family aspects.<sup>51</sup> The scale was designed to assess the QoL of healthy individuals and people with an illness. All items are measured using a 6-point Likert scale. The original scale was validated in 48 studies and in 13 different types of diseases. The internal consistency reliability across those studies ranged from 0.73 to 0.99.<sup>52</sup>

The review found only one translated study that was culturally adapted and which tested the reliability and content validity of OLI in modern standard Arabic.<sup>53</sup> The study involved healthy individuals, and hypertensive, diabetic, cancer, and dialysis patients. The reliability was

adequate, with coefficients ranging from 0.88 to 0.97. The Arabic version of the QLI met all the quality criteria with regard to cross-cultural adaptation.

### **Discussion:**

This review identified only seven QoL scales that were translated and tested for validity and reliability in the Arabic-speaking population. No scales were initially developed specifically for the Arabic-speaking population as most of the QoL scales were developed for use within the English language. There is consistency among the QoL scales as they generally assess physical, psychological, and other important aspects of life. To understand QoL in Arabic-speaking nations, there is a great need for more translated and culturally-adapted QoL scales. This is because only two studies included in this review were adapted along the lines of all five steps of the Guillemin *et al.* guidelines.

None of the studies in this review evaluated all nine quality assessment criteria. As a result, further psychometric studies are required to improve the validity and reliability of the Arabic version of the QoL scales. Prior to use in an Arabic-speaking population, the reliability and validity of the QoL scale should be evaluated.

The selection of QoL scales is dependent on a number of different factors such as the demographic and characteristic variables of the sample, the psychometric properties of the scale, and the number of scale items. Most importantly, researchers need to consider the aspects of life that require evaluation, to facilitate the choice of sub-scale dimensions in their appropriate population of interest. For example, if the sample used consisted of Arab cancer patients, the FACT-G or the EORTC QLQ-C30 would provide cancer-specific QoL information. Both scales demonstrated acceptable psychometric properties among Arab cancer patients, and have been tested psychometrically in the general population using other languages.<sup>54,55</sup> If the sample consisted of the general Arabic-speaking population, the SF-36 might be a better choice to provide general QoL information.

The studies in this review conducted psychometric testing in a variety of populations. The general population was the most frequently studied, perhaps because this provides a large sample size which improves the psychometric evaluation. The scales contain a range of items (from 5 to 70). Six scales contained fewer than 36 items, and could be administered in between five and ten minutes. Only the QLI had 70 items, and the administration time was reported as being approximately 10 minutes. Unlike other QoL scales, the QLI weights satisfaction in a

particular aspect of life in terms of importance, so that items with high satisfaction and importance scores are scored the highest, while items with low satisfaction and low importance receive the lowest values. The questionnaire length could be an obstacle when conducting research in a clinical setting.<sup>56</sup> Reducing the number of items in the questionnaire could therefore increase the response rate.<sup>57</sup>

The cross-cultural adaptation procedure with regard to the scale has a potential effect on the scale's credibility. Consequently, cultural adaptation is required to ensure that the content of the translated scale is equivalent to the original, using a systematic and standardised approach. Guillemin *et al.*<sup>12</sup> proposed guidelines that appear to offer an appropriate approach, and proposed steps to follow for translation and cultural adaptation purposes. Adhering to the guidelines produces cultural equivalence and maximises the acceptability of the linguistic structure of the QoL scale. After adopting those steps in this review, most of the scales did not provide sufficient information with regard to the translation and adaptation processes. Therefore, it is recommended that in future studies researchers identify the sequential stages of cross-cultural adaptation of the scale by recognising the elements of the guidelines at each stage.

This review has a number of limitations. It focuses on the psychometric properties only. Further reviews could evaluate the methodological quality of the studies by using one of the guidelines. One author only screened and assessed the eligibility of the papers. This may induce the risk of bias or possible errors in data collection. Although this review includes 27 studies using different Medical Subject Headings (MeSH), it is possible that some relevant studies were not included.

## **Conclusion:**

The review aims to evaluate the psychometric properties and cultural adaptation process of Arabic QoL scales to support current and future use of the scales based on evidence. However, the selection of the QoL scale will be influenced by the objective of the study, the intervention, and the domains of the scales. This review provides empirical evidence regarding the psychometric testing of QoL scales that have been used in studies of the Arabic-speaking population. In general, the scales reviewed provided insufficient information about the extent of cultural adaptation. Most scales provided information regarding content, construct validity and internal consistency, while information related to agreement, responsiveness, floor and ceiling effects and interpretation was lacking. Specifically, test-retest reliability, criterion

validity and the sensitivity of the Arabic QoL scales require further testing. Furthermore, future translations and cultural adaptations of Arabic QoL scales should use the available guidelines such as Guillemin recommends to ensure the quality of the scales.

## References

- 1 Romero M, Vivas-Consuelo D, Alvis-Guzman N. Is Health Related Quality of Life (HRQoL) a valid indicator for health systems evaluation? SpringerPlus 2013; 2: 664.
- 2 Verdugo MA, Schalock RL, Keith KD, Stancliffe RJ. Quality of life and its measurement: important principles and guidelines. J Intellect Disabil Res 2005; 49: 707–717.
- 3 Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000; 25: 3186–3191.
- 4 IstiZada. Arabic Speaking Countries List - 2018 | IstiZada. 2018.<http://istizada.com/complete-list-of-arabic-speaking-countries-2014/> (accessed 13 May2018).
- 5 Sadat F, Kazemi F, Farzindar A. Automatic Identification of Arabic Dialects in Social Media. In: Proceedings of the First International Workshop on Social Media Retrieval and Analysis. ACM: New York, NY, USA, 2014, pp 35–40.
- 6 Coons SJ, Alabdulmohsin SA, Draugalis JR, Hays RD. Reliability of an Arabic Version of the RAND-36 Health Survey and Its Equivalence to the US-English Version. Med Care 1998; 36: 428–432.
- 7 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010; 8: 336–341.
- 8 Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007; 60: 34–42.
- 9 Downing SM. Reliability: on the reproducibility of assessment data. Med Educ 2004; 38: 1006–1012.
- 10 Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med 2006; 119: 166–e7.
- 11 Messick S. Standards of Validity and the Validity of Standards in Performance Assessment. Educ Meas 1995; 14: 5–8.
- 12 Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. J Clin Epidemiol 1993; 46: 1417–1432.
- 13 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.
- 14 Touma Z, Ghandour L, Sibai A, Puzantian H, Hamdan A, Hamdan O et al. Cross-cultural adaptation and validation of Behçet’s disease quality of life questionnaire. BMC Med Res Methodol 2011; 11. doi:10.1186/1471-2288-11-52.

- 15 Shata ZN, Hm E-K, Ibrahim H. Reliability and Validity of an Arabic Version of Quality of life - Alzheimer Disease in Alexandria, Egypt. *Int J Behav Res Psychol* 2015; : 157–163.
- 16 Zahid MA, Ohaeri JU, Elshazly AS, Basiouny MA, Hamoda HM, Varghese R. Correlates of quality of life in an Arab schizophrenia sample. *Soc Psychiatry Psychiatr Epidemiol* 2010; 45: 875–887.
- 17 Adnane C, Oubahmane T, Adouly T, Elhani L, Rouadi S, Abada RL et al. Cross-cultural and Moroccan Validation of the University of Washington Quality of Life Questionnaire for Patients With Head and Neck Cancer. *Ann Otol Rhinol Laryngol* 2016; 125: 151–159.
- 18 Ouattassi N, Benmansour N, ElFakir S, Nejari C, Alami MN. Translation and validation of EORTC QLQ-H&N 35 into Moroccan Arabic for ENT head and neck cancer patients in Morocco. *Eur Arch Otorhinolaryngol* 2016; 273: 2727–2734.
- 19 El Fakir S, Abda N, Bendahhou K, Zidouh A, Bennani M, Errihani H et al. The European Organization for Research and Treatment of Cancer quality of life questionnaire-BR23 Breast Cancer-Specific Quality of Life Questionnaire: psychometric properties in a Moroccan sample of breast cancer patients. *BMC Res Notes* 2014; 7: 53.
- 20 Hoopman R, Terwee CB, Devillé W, Knol DL, Aaronson NK. Evaluation of the psychometric properties of the SF-36 health survey for use among Turkish and Moroccan ethnic minority populations in the Netherlands. *Qual Life Res* 2009; 18: 753–764.
- 21 Hoopman R, Terwee CB, Aaronson NK. Translated COOP/WONCA charts found appropriate for use among Turkish and Moroccan ethnic minority cancer patients. *J Clin Epidemiol* 2008; 61: 1036–1048.
- 22 Hoopman R, Terwee CB, Muller MJ, Aaronson NK. Translation and validation of the SF-36 Health Survey for use among Turkish and Moroccan ethnic minority cancer patients in The Netherlands. *Eur J Cancer* 2006; 42: 2982–2990.
- 23 Guermazi M, Allouch C, Yahia M, Huissa TBA, Ghorbel S, Damak J et al. Translation in Arabic, adaptation and validation of the SF-36 Health Survey for use in Tunisia. *Ann Phys Rehabil Med* 2012; 55: 388–403.
- 24 Younsi M, Chakroun M. Measuring health-related quality of life: psychometric evaluation of the Tunisian version of the SF-12 health survey. *Qual Life Res* 2014; 23: 2047–2054.
- 25 El-Kalla R, Khalaf M, Saad M, Othman E. Reliability of the Arabic Egyptian Version of Short Form 36 Health Survey Questionnaire to Measure Quality of Life in Burned Patient. *Med J Cairo Univ* 2016; 84: 311–316.
- 26 Polit DF, Beck CT. *Nursing Research: Generating and Assessing Evidence for Nursing Practice*. 10th ed. Lippincott Williams & Wilkins: Philadelphia, 2017.
- 27 Sabbah I, Drouby N, Sabbah S, Retel-Rude N, Mercier M. Quality of Life in rural and urban populations in Lebanon using SF-36 Health Survey. *Health Qual Life Outcomes* 2003; 1: 30.

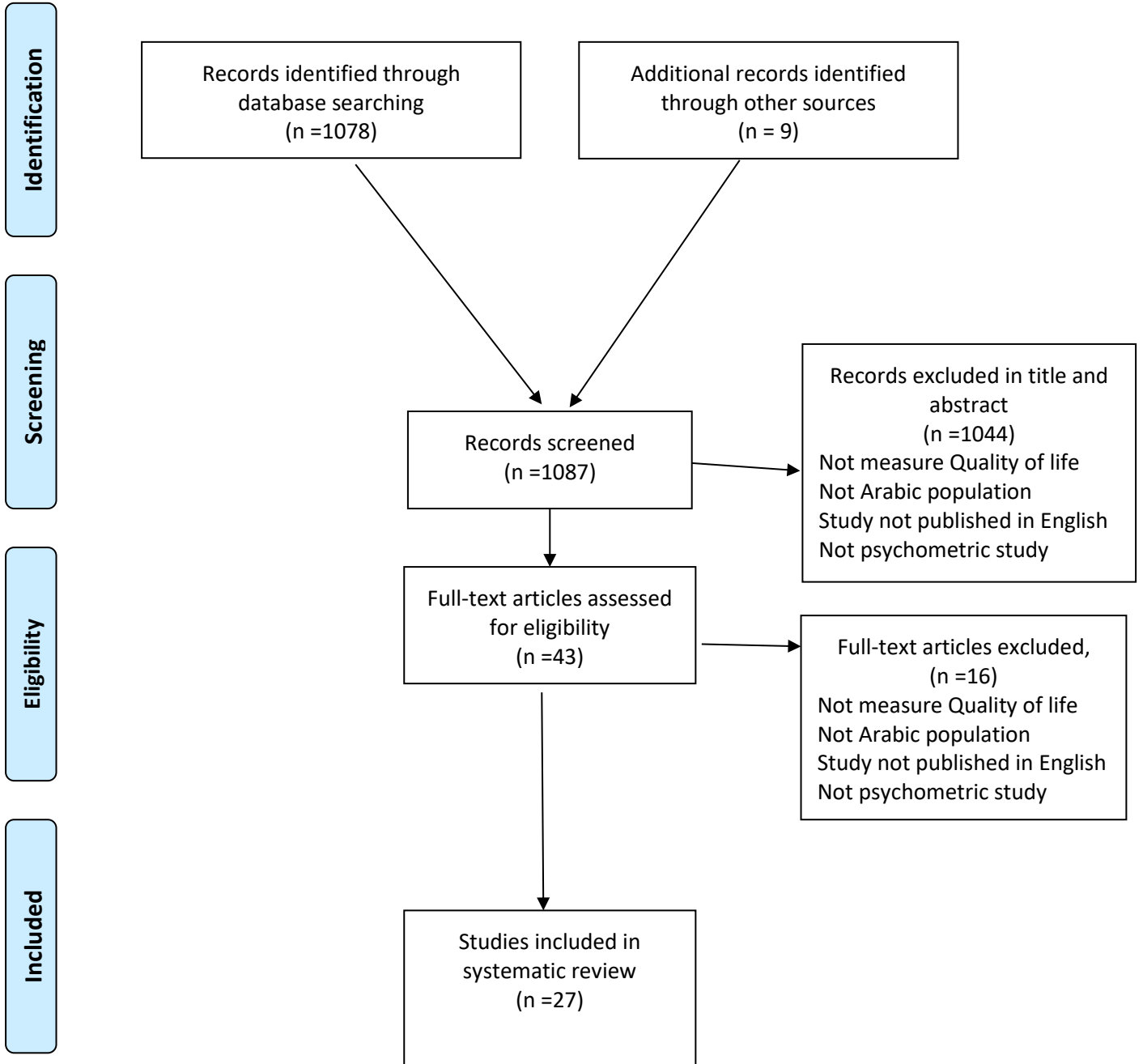
- 28 Dalky HF, Meininger JC, Al-Ali NM. The Reliability and Validity of the Arabic World Health Organization Quality of Life-BREF Instrument Among Family Caregivers of Relatives With Psychiatric Illnesses in Jordan. *J Nurs Res* 2017; 25: 224–230.
- 29 Ohaeri JU, Awadalla AW, El-Abassi A-HM, Jacob A. Confirmatory factor analytical study of the WHOQOL-Bref: experience with Sudanese general population and psychiatric samples. *BMC Med Res Methodol* 2007; 7: 37.
- 30 Ohaeri JU, Awadalla AW. The reliability and validity of the short version of the WHO Quality of Life Instrument in an Arab general population. *Ann Saudi Med* 2009; 29: 98–104.
- 31 Khader S, Hourani MM, Al-Akour N. Normative data and psychometric properties of short form 36 health survey (SF-36, version 1.0) in the population of north Jordan. *East Mediterr Health J* 2011; 17: 368–374.
- 32 Soudy H, Maghfour I, Elhassan TAM, Abdullah E, Rauf SM, Al Zahrani A et al. Translation and validation of the Functional Assessment of Cancer Therapy-Bone Marrow Transplant (FACT-BMT) version 4 quality of life instrument into Arabic language. *Health Qual Life Outcomes* 2018; 16: 47.
- 33 Al Barmawi M, Al Hadid LA, Alqudah HN, Al Hadid WA, Shamoun SA. Measuring the Quality of Life among Head-and/or-Neck Cancer Patients with Oral Mucositis Using the Functional Assessment of Cancer Therapy-General in Jordan. *Asia-Pac J Oncol Nurs* 2018; 5: 320–326.
- 34 Alawneh A, Yasin H, Khirfan G, Qayas BA, Ammar K, Rimawi D et al. Psychometric properties of the Arabic version of EORTC QLQ-C15-PAL among cancer patients in Jordan. *Support Care Cancer* 2016; 24: 2455–2462.
- 35 Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *J Natl Cancer Inst* 1993; 85: 365–376.
- 36 Awad MA, Denic S, El Taji H. Validation of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaires for Arabic-speaking Populations. *Ann N Y Acad Sci* 2008; 1138: 146–154.
- 37 Bener A, Alsulaiman R, Doodson L, El Ayoubi HR. An assessment of reliability and validity of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire C30 among breast cancer patients in Qatar. *J Fam Med Prim Care* 2017; 6: 824–831.
- 38 Huijter HA-S, Sagherian K, Tamim H. Validation of the Arabic version of the EORTC quality of life questionnaire among cancer patients in Lebanon. *Qual Life Res* 2013; 22: 1473–1481.
- 39 Groenvold M, Petersen MA, Aaronson NK, Arraras JI, Blazeby JM, Bottomley A et al. The development of the EORTC QLQ-C15-PAL: a shortened questionnaire for cancer patients in palliative care. *Eur J Cancer* 2006; 42: 55–64.



- 40 Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *J Clin Oncol* 1993; **11**: 570–579.
- 41 Lazenby M, Khatib J, Al-Khair F, Neamat M. Psychometric properties of the Functional Assessment of Chronic Illness Therapy--Spiritual Well-being (FACIT-Sp) in an Arabic-speaking, predominantly Muslim population. *Psychooncology* 2013; **22**: 220–227.
- 42 Zahran MH, Sheir K, Zidan EM, Ali-El-Dein B. Validation of the Arabic version of the Functional Assessment of Cancer Therapy-Bladder questionnaire in Egyptian patients with bladder cancer. *Arab J Urol* 2017; **15**: 110–114.
- 43 McHorney CA, Ware JE, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994; **32**: 40–66.
- 44 Khoudri I, Ali Zeggwagh A, Abidi K, Madani N, Abouqal R. Measurement properties of the short form 36 and health-related quality of life after intensive care in Morocco. *Acta Anaesthesiol Scand* 2007; **51**: 189–197.
- 45 Sheikh KA, Yagoub U, Elsatouhy M, Al Sanosi R, Mohamud SA. Reliability and Validity of the Arabic Version of the SF-36 Health Survey Questionnaire in Population of Khat Chewers—Jazan Region-Kingdom of Saudi Arabia. *Appl Res Qual Life* 2015; **10**: 1–13.
- 46 Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001; **33**: 337–343.
- 47 Bekairy AM, Bustami RT, Almotairi M, Jarab A, Katheri AM, Aldebasi TM *et al.* Validity and reliability of the Arabic version of the the EuroQOL (EQ-5D). A study from Saudi Arabia. *Int J Health Sci* 2018; **12**: 16–20.
- 48 WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. *Psychol Med* 1998; **28**: 551–558.
- 49 Bani-Issa W. Evaluation of the health-related quality of life of Emirati people with diabetes: integration of sociodemographic and disease-related variables. *East Mediterr Health J* 2011; **17**: 825–830.
- 50 Stewart AL, Ware JE. *Measuring Functioning and Well-being: The Medical Outcomes Study Approach*. Duke University Press, 1992.
- 51 Ferrans CE, Powers MJ. Quality of life index: development and psychometric properties. *Adv Nurs Sci* 1985; **8**: 15–24.
- 52 Ferrans CE, Powers MJ. QLI Reliability and Validity. QLI Reliab. Validity. 2018.<https://qli.org.uic.edu/reliability/reliabilityhome.htm> (accessed 1 Nov2018).
- 53 Halabi JO. Psychometric properties of the Arabic version of Quality of Life Index. *J Adv Nurs* 2006; **55**: 604–610.

- 54 Hinz A, Glaesmer H, Brähler E, Löffler M, Engel C, Enzenbach C *et al.* Sleep quality in the general population: psychometric properties of the Pittsburgh Sleep Quality Index, derived from a German community sample of 9284 people. *Sleep Med* 2017; **30**: 57–63.
- 55 Pearman T, Yanez B, Peipert J, Wortman K, Beaumont J, Cella D. Ambulatory cancer and US general population reference values and cutoff scores for the functional assessment of cancer therapy. *Cancer* 2014; **120**: 2902–2909.
- 56 Mark TL, Johnson G, Fortner B, Ryan K. The benefits and challenges of using computer-assisted symptom assessments in oncology clinics: results of a qualitative assessment. *Technol Cancer Res Treat* 2008; **7**: 401–406.
- 57 Rolstad S, Adler J, Rydén A. Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value Health* 2011; **14**: 1101–1108.
- 58 Awad MA, Denic S, El H. Validation of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaires for Arabic-speaking Populations. 2008; **154**: 146–154.
- 59 Alawadhi SA, Ohaeri JU. Validity and reliability of the European Organization for Research and Treatment in Cancer Quality of Life Questionnaire (EORTC QLQ): experience from Kuwait using a sample of women with breast cancer. *Ann Saudi Med* 2010; **30**: 390–396.
- 60 Hoopman R, Terwee CB, Devillé W, Knol DL, Aaronson NK. Evaluation of the psychometric properties of the SF-36 health survey for use among Turkish and Moroccan ethnic minority populations in the Netherlands. *Qual Life Res* 2009; **18**: 753–764.
- 61 Khoudri I, Ali Zeggwagh A, Abidi K, Madani N, Abouqal R. Measurement properties of the Short Form 36 and health-related quality of life after intensive care in Morocco. *Acta Anaesthesiol Scand* 2007; **51**: 189–197.
- 62 Guermazi M, Allouch C, Yahia M, Huissa TBA, Ghorbel S, Damak J *et al.* Translation in Arabic, adaptation and validation of the SF-36 Health Survey for use in Tunisia. *Ann Phys Rehabil Med* 2012; **55**: 388–403.

**Figure 1: PRISMA diagram QoL Scales**



**Table 1: Quality Criteria for Measurement Properties of Scales <sup>c</sup>**

	Property	Definition		Quality Criteria <sup>a,b</sup>
1	Content validity	The amount to which the domain of Interest is comprehensively sampled by the items in the questionnaire	+	A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection;
			?	A clear description of above-mentioned aspects is missing OR only target Population involved OR doubtful design or method;
			-	No target population involvement;
			0	No information found on target population involvement.
2	Internal Consistency	The amount to which items in a (sub) scale Are intercorrelated, so measuring the same construct	+	Factor analyses performed on adequate sample size (7 * # items and > 100) AND Cronbach's alpha (s) calculated per dimension AND Cronbach's alpha (s) Between 0.70 and 0.95;
			?	No factor analysis OR doubtful design or method;
			-	Cronbach's alpha (s) < 0.70 or > 0.95, despite adequate design and method;
			0	No information found on internal consistency.
3	Criterion validity	The extent to which scores on a Particular questionnaire refer to a gold Standard	+	Convincing arguments that gold standard is " gold " AND correlation with gold standard > 0.70;
			?	No convincing arguments that gold standard is " gold " OR doubtful design or Method;
			-	Correlation with gold standard < 0.70, continuous adequate design and method;
			0	No information found on criterion validity.
4	Construct validity	The amount to which scores on a Particular questionnaire refer to other Measures in a manner that is consistent with theoretically derived hypotheses Relating the concepts that are being measured	+	Specific hypotheses were formed and at least 75% of the results are in accordance with these hypotheses;
			?	Doubtful design or method (e.g. no hypotheses)
			-	Less than 75% of hypotheses were confirmed, despite adequate design and Methods;
			0	No information found on construct validation.
5	Reproducibility			
	5.1. Agreement	The amount to which the scores on repeated measures are close to	+	MIC! SDC OR MIC outside the LOA OR convicting arguments that agreement is acceptable;

		each other (absolute measurement error)	?	Doubtful design or method OR (MIC not defined AND no convincing arguments that agreement is acceptable)
			-	MIC > SDC OR MIC equals or inside LOA, despite adequate design and method;
			0	No information found on agreement.
	5.2. Reliability	The amount to which patients can be Distinguished from each other, despite Measurement errors (Relative measurement error)	+	ICC or weighted Kappa > 0.70;
			?	Doubtful design or method (e.g., time interval not mentioned);
			-	ICC or weighed Kappa! 0.70, despite adequate design and method;
			0	No information found on reliability.
6	Responsive ness	The ability of a questionnaire to detect Clinically important changes over time	+	SDC or SDC! MIC OR MIC outside the LOA OR RRO1.96 OR AUC > 0.70;
			?	Doubtful design or method;
			-	SDC or SDC > MIC OR MIC equals or inside LOA OR RR < 1.96 OR AUC! 0.70, despite adequate design and methods
			0	No information found on responsiveness.
7	Floor and ceiling Effects	The number of responders who achieved the lowest or highest possible score	+	<15% of the respondents achieved the highest or lowest possible scores;
			?	Doubtful design or method;
			-	>15% of the respondents achieved the highest or lowest possible scores, strict adequate design and methods;
			0	No information found on interpretation.
8	Interpretability	The degree to which one can assign Qualitative meaning to quantitative scores	+	Mean and SD scores presented at least four relevant subgroups of patients and MIC defined;
			?	Doubtful design or method OR less than four subgroups OR no MIC defined;
			0	No information found on interpretation.
<p>MIC= minimal important change; SDC=smallest detectable change; LOA=limits of agreement; ICC=Intraclass correlation; SD, standard deviation.  <sup>a</sup> + = positive rating; ?=indeterminate rating; - =negative rating; 0=no information available.  <sup>b</sup> Doubtful design or method= lacking of a clear description of the design or methods of the study, sample size smaller than 50 subjects (should be at least 50 in every (subgroup) analysis), or any important methodological weakness in the design or execution of the study.</p>				

<sup>c</sup> This table adapted from Terwee et al. 2007 page 39 <sup>8</sup>

**Table 2: Characteristics of the Studies**

	Scale with references	Country	Design	Type of participants	Sample Size	Language	Internal consistency Cronbach's Coefficient Alpha
1	EORTC QLQ-C30 Huijer <i>et al.</i> <sup>38</sup>	Lebanon	Cross-Sectional	Mixed cancer patients	200	Standard Arabic	Overall C. $\alpha > 0.70$ C. $\alpha$ 0.38 to 0.80
2	EORTC QLQ-C30 Awad <i>et al.</i> <sup>36</sup>	UAE	Cross-Sectional	Breast cancer	87	Standard Arabic	Overall C. $\alpha > 0.70$ C. $\alpha$ 0.51 to 0.84
3	EORTC QLQ-C30 Alawadhi and Ohaeri <sup>59</sup>	Kuwait	Cross-Sectional	Breast cancer	348	Standard Arabic	Overall C. $\alpha = 91$ C. $\alpha$ 0.51 to 0.84
4	EORTC QLQ-C30 Bener <i>et al.</i> <sup>37</sup>	Qatar	Cross-Sectional	Breast cancer	678	Standard Arabic	Overall C. $\alpha = 91$ C. $\alpha$ 0.55 to 0.89
5	EORTC QLQ-C15-PAL Alawneh <i>et al.</i> <sup>34</sup>	Jordan	Cross-Sectional	Mixed cancer patients	175	Standard Arabic	Overall C. $\alpha > 0.70$ C. $\alpha$ 0.72 to 0.90
6	FACT-G Lazenby <i>et al.</i> <sup>41</sup>	Jordan	Cross-Sectional	Mixed cancer patients	205	Standard Arabic	C. $\alpha$ 0.80 to 0.83
7	FACT-G Zahran <i>et al.</i> <sup>42</sup>	Egyptian	Cross-Sectional	Bladder cancer	90	Standard Arabic	FACT- G 0.80-0.94
8	FACT-G Al Barmawi <i>et al.</i> <sup>33</sup>	Jordan	Cross-Sectional	head-and / or-neck cancer patients	118	Standard Arabic	Overall C. $\alpha = 76$ C. $\alpha$ 0.67 to 0.83
9	FACT-G Soudy <i>et al.</i> <sup>32</sup>	Saudi Arabia	Cross-Sectional	Lymphoma patients	108	Standard Arabic	Overall C. $\alpha = 89$ C. $\alpha$ 0.67 to 0.88

				underwent stem cell transplant			
10	RAND 36-Item OR SF-36 Coons et al. <sup>6</sup>	Saudi Arabia	Longitudinal	General population	415	Standard Arabic	C. $\alpha$ 0.60 to 0.87
11	SF-36 Sabbah et al. <sup>27</sup>	Lebanon	Cross-Sectional	General population	524	Standard Arabic	C. $\alpha$ 0.70 to 0.90
12	SF-36 Hoopman et al. <sup>60</sup>	Netherlands	Longitudinal	General population	Turkish (N = 409) Moroccan (N = 377) Dutch (N = 9,628)	Tarifit (Local dialect)	C. $\alpha$ 0.63 to 0.93
13	SF-36 Hoopman et al. <sup>22</sup>	Netherlands	Longitudinal	Mixed cancer patients	79 Moroccan patients 90 Turkish	Tarifit (Local dialect)	C. $\alpha$ 0.65 to 0.94
14	SF-36 Khoudri et al. <sup>61</sup>	Morocco	Cross-Sectional	Mixed patients discharged from ICU	145	Standard Arabic	Overall C. $\alpha \geq .70$ C. $\alpha$ 0.84 to 0.99
15	SF-36 Guermazi et al. <sup>62</sup>	Tunisia	Cross-Sectional	General population	130	Tunisian (Local dialect)	Overall C. $\alpha = 0.94$ C. $\alpha$ 0.72 to 0.89
16	SF-36 El-Kalla et al. <sup>25</sup>	Egypt	Longitudinal	Patient with burn injury	40	Egyptian (Local dialect)	Overall C. $\alpha = 0.8$

17	SF-36 Sheikh et al. <sup>45</sup>	Saudi Arabia	Cross-Sectional	Khat Chewers	300	Standard Arabic	Overall C. $\alpha = 0.94$ C. $\alpha 0.72$ to $0.90$
18	SF-36 Khader et al. <sup>31</sup>	Jordan	Cross-Sectional	General population	511	Standard Arabic	C. $\alpha 0.71$ to $0.90$
19	SF-12 Younsi and Chakroun <sup>24</sup>	Tunisia	Cross-Sectional	General population	3582	Tunisian (Local dialect)	C. $\alpha 0.73$
20	EQ-5D Aburuz et al. <sup>13</sup>	Jordan	Cross-Sectional	General population	186	Standard Arabic	Overall C. $\alpha \geq .75$
21	EQ-5D Bekairy et al. <sup>47</sup>	Saudi Arabia	Longitudinal	Mixed Patients	80	Standard Arabic	Overall C. $\alpha = .72$
22	WHOQOL-Bref Ohaeri and Awadalla <sup>30</sup>	Kuwait	Longitudinal	General population	3303	Standard Arabic	Overall C. $\alpha \geq 0.90$ C. $\alpha 0.69$ to $0.83$
23	WHOQOL-Bref Ohaeri et al. <sup>29</sup>	Sudan	Cross-Sectional	General population, Psychiatric patients	623	Standard Arabic	General population C. $\alpha = 0.88$ , Psychiatric patients C. $\alpha = 0.93$ and Caregivers C. $\alpha = 0.92$
24	WHOQOL-Bref Bani-Issa <sup>49</sup>	UAE	Cross-Sectional	Diabetic patients	200	Standard Arabic	Overall C. $\alpha = 0.85$ C. $\alpha 0.89$ to $0.91$
25	WHOQOL-Bref Dalky et al. <sup>28</sup>	Jordan	Cross-Sectional	Family caregiver of patient	266	Standard Arabic	Overall C. $\alpha = 0.92$
26	COOP/WONCA Hoopman et al. <sup>21</sup>	Moroccan	Cross-Sectional	Mixed cancer patients	37 Arabic 97 Turkish	Moroccan Tarifit (Local dialect)	IC not reported



27	Quality of Life Index (QLI) Halabi <sup>53</sup>	Jordan	Longitudinal	General population	35	Standard Arabic	Overall C. $\alpha = 0.90$

**Table 3: The Assessment of Measurement Properties of Quality of Life Scales**

	Scale	Content Validity	Internal Consistency	Criterion Validity	Construct Validity	Reproducibility		Responsiveness	Floor and Ceiling Effects	Interpretation
						Agreement	Reliability			
1	EORTC QLQ-C30 Huijer <i>et al.</i> <sup>38</sup>	+	+	0	+	0	0	0	0	0
2	EORTC QLQ-C30 Awad <i>et al.</i> <sup>36</sup>	+	+	0	+	0	0	0	0	0
3	EORTC QLQ-C30 Alawadhi and Ohaeri <sup>59</sup>	+	+	0	0	0	0	0	0	0
4	EORTC QLQ-C30 Bener <i>et al.</i> <sup>37</sup>	+	+	0	+	0	0	0	0	0
5	EORTC QLQ-C15-PAL Alawneh <i>et al.</i> <sup>34</sup>	+	+	0	+	0	+	0	0	0
6	FACT-G Lazenby <i>et al.</i> <sup>41</sup>	+	+	0	+	0	0	0	0	0
7	FACT-G Zahran <i>et al.</i> <sup>42</sup>	+	+	0	+	0	0	0	0	0
8	FACT-G Al Barmawi <i>et al.</i> <sup>33</sup>	+	+	0	+	0	0	0	0	0
9	FACT-G Soudy <i>et al.</i> <sup>32</sup>	+	+	0	+	0	0	0	0	0
10	RAND 36-Item OR SF-36 Coons <i>et al.</i> <sup>6</sup>	+	+	0	+	0	+	0	0	0
11	SF-36 Sabbah <i>et al.</i> <sup>27</sup>	+	+	+	+	0	0	0	+	0
12	SF-36 Hoopman <i>et al.</i> <sup>60</sup>	+	+	0	0	0	0	0	+	0
13	SF-36 Hoopman <i>et al.</i> <sup>22</sup>	+	+	0	+	0	0	+	+	0
14	SF-36 Khoudri <i>et al.</i> <sup>61</sup>	+	+	0	+	0	+	0	0	0
15	SF-36 Guermazi <i>et al.</i> <sup>62</sup>	+	+	0	+	+	+	0	0	0
16	SF-36 El-Kalla <i>et al.</i> <sup>25</sup>	0	+	0	0	0	+	0	0	0
17	SF-36 Sheikh <i>et al.</i> <sup>45</sup>	0	+	0	+	0	0	0	0	0
18	SF-36 Khader <i>et al.</i> <sup>31</sup>	0	+	0	+	0	0	0	+	0
19	SF-12 Younsi and Chakroun <sup>24</sup>	0	+	0	-	0	0	0	0	0
20	EQ-5D Aburuz <i>et al.</i> <sup>13</sup>	+	+	0	+	0	+	0	0	-
21	EQ-5D Bekairy <i>et al.</i> <sup>47</sup>	+	+	0	+	0	+	0	0	0
22	WHOQOL-Bref Ohaeri and Awadalla <sup>30</sup>	+	+	0	+	0	-	0	-	0
23	WHOQOL-Bref Ohaeri <i>et al.</i> <sup>29</sup>	+	+	0	+	0	-	0	0	0
24	WHOQOL-Bref Bani-Issa <sup>49</sup>	0	+	0	0	0	0	0	0	0
25	WHOQOL-Bref Dalky <i>et al.</i> <sup>28</sup>	+	+	+	+	0	0	0	0	0
26	COOP/WONCA Hoopman <i>et al.</i> <sup>21</sup>	+	0	0	+	0	0	-	-	0
27	Quality of Life Index (QLI) Halabi <sup>53</sup>	+	+	0	0	0	0	0	0	0

Rating: + = positive, 0= intermediate, - = negative, ? = information not reported

**Table 4: Cross-Cultural Adaptation of Quality of Life Scales**

	Study and Scale	Translation Technique	Back-Translation Technique	Committee Approach	Pre-testing	Weighting Score Adaptation	Overall Score
1	EORTC QLQ-C30 Huijer <i>et al.</i> <sup>38</sup>	3	3	3	3	NA	3
2	EORTC QLQ-C30 Awad <i>et al.</i> <sup>36</sup>	NA	NA	NA	NA	NA	NA
3	EORTC QLQ-C30 Alawadhi and Ohaeri <sup>59</sup>	NA	NA	NA	NA	NA	NA
4	EORTC QLQ-C30 Bener <i>et al.</i> <sup>37</sup>	NA	NA	NA	NA	NA	NA
5	EORTC QLQ-C15-PAL Alawneh <i>et al.</i> <sup>34</sup>	2	NR	NR	NR	NR	2
6	FACT-G Lazenby <i>et al.</i> <sup>41</sup>	3	3	3	3	NA	3
7	FACT-G Zahran <i>et al.</i> <sup>42</sup>	3	3	3	3	NA	3
8	FACT-G Al Barmawi <i>et al.</i> <sup>33</sup>	3	3	NR	3	NR	3
9	FACT-G Soudy <i>et al.</i> <sup>32</sup>	3	3	3	3	3	3
10	RAND 36-Item OR SF-36 Coons <i>et al.</i> <sup>6</sup>	3	3	2	3	3	2.8
11	SF-36 Sabbah <i>et al.</i> <sup>27</sup>	3	3	3	3	NR	3
12	SF-36 Hoopman <i>et al.</i> <sup>60</sup>	NA	NA	NA	NA	NA	NA
13	SF-36 Hoopman <i>et al.</i> <sup>22</sup>	2	1	NR	NR	NR	1.5
14	SF-36 Khoudri <i>et al.</i> <sup>61</sup>	NA	NA	NA	NA	NA	NA
15	SF-36 Guermazi <i>et al.</i> <sup>62</sup>	3	3	1	3	NR	2.5
16	SF-36 El-Kalla <i>et al.</i> <sup>25</sup>	3	3	3	3	NR	3
17	SF-36 Sheikh <i>et al.</i> <sup>45</sup>	NA	NA	NA	NA	NA	NA
18	SF-36 Khader <i>et al.</i> <sup>31</sup>	NA	NA	NA	NA	NA	NA
19	SF-12 Younsi and Chakroun <sup>24</sup>	NA	NA	NA	NA	NA	NA
20	EQ-5D Aburuz <i>et al.</i> <sup>13</sup>	3	2	2	3	NA	2.5
21	EQ-5D Bekairy <i>et al.</i> <sup>47</sup>	NA	NA	NA	NA	NA	NA
22	WHOQOL-Bref Ohaeri and Awadalla <sup>30</sup>	NA	NA	NA	NA	NA	NA
23	WHOQOL-Bref Ohaeri <i>et al.</i> <sup>29</sup>	NA	NA	NA	NA	NA	NA
24	WHOQOL-Bref Bani-Issa <sup>49</sup>	NA	NA	NA	NA	NA	NA
25	WHOQOL-Bref Dalky <i>et al.</i> <sup>28</sup>	NA	NA	NA	NA	NA	NA

26	COOP/WONCA Hoopman et al. <sup>21</sup>	2	2	3	3	NR	2.5
27	Quality of Life Index (QLI) Halabi <sup>53</sup>	3	3	3	3	NR	3

Rating: 3 = Good, 2 = moderate, 1 = poor, NR = information not reported, NA: not applicable