

# Comparison of Instruments for Measuring Health-Related Quality of Life and Their Applicability to Ehealth Applications

Marlene SCHMIDT<sup>a</sup> and Martin STAEMMLER<sup>1a</sup>

<sup>a</sup> *University of Applied Sciences, Stralsund, Germany*

**Abstract.** Health-related Quality of Life (HRQoL) assessment has proven as a good means for assessing treatment options or impact of applications supporting the patient in adherence, monitoring and better understanding of health issues. While most of the HRQoL instruments were designed several years ago, their capability to assess the impact of ehealth application is in question. The objective of this paper is to assess HRQoL instruments including a focus on the ehealth domain. Methods: Generic and specific instruments are selected based on their widespread use. Published criteria for assessing HRQoL instruments are used for a baseline, which are amended by criteria covering both the ehealth domain and the conditions of use of instruments and structured in groups. Results: Seven instruments have been selected and assessed using the established criteria. The instruments scored differently regarding the ehealth domain, however overall rather low. Applying weighting per group allows highlighting specific aspects. Based on the assessment, further research should consider the development of a ehealth domain module as part of the specific instruments.

**Keywords.** Health-Related Quality of Life, ehealth applications

## 1. Introduction

Nowadays Health-related Quality of Life (HRQoL) assessments are widely used for informed decisions on treatment options and determination of the outcome of care [1-4]. A variety of tested and standardized HRQoL instruments exist, as listed in [5-7]. Some of them aim for generic use, for example Short-Form 36 Health Survey (SF-36), Nottingham Health Profile (NHP), WHO instruments or the EQ-5D-5L [8-11]. However, they represent different viewpoints in defining and understanding HRQoL, e.g. with more or less focus on living context. Others are disease oriented or allow adding disease specific modules such as the EORTC QLQ-C30 [12] or the Functional assessment of Chronic Illness Therapy (FACIT-G) [13]. This orientation also applies to most published comparisons of HRQoL instruments [14-20] which start from a disease and conduct an analysis or meta search in literature based on the most appropriate or widely used HRQoL instrument. Consequently the criteria used appear to be biased by the targeted disease. Furthermore, most of the HRQoL instruments have been developed several

---

<sup>1</sup> Corresponding Author, Martin Staemmler, Health Informatics, University of Applied Sciences, Zur Schwedenschanze 15, 18435 Stralsund, Germany; E-mail: martin.staemmler@fh-stralsund.de.

years ago, thereby having limited focus on assessing the impact of today's ehealth applications on HRQoL.

The objective of this paper is to establish criteria for the comparison of HRQoL instruments with a particular focus on ehealth applications, to apply the criteria for a comparison of selected instruments and to derive a recommendation for assessing HRQoL resulting when using ehealth applications.

## 2. Methods

HRQoL instruments can be classified according to their scope. While disease specific instruments target both the most prevalent (e.g. coronary, pulmonary, cancer, diabetic) and specific (e.g. gestational) diseases, the generic ones are applicable in an overarching way. Furthermore, they differ in the result. Index oriented instruments provide one value, representing the level of HRQoL measured, whereas profile oriented instruments reveal an assessment for each dimensions used in the assessment. The number of dimensions reflects the understanding of HRQoL, e.g. addition of a behavioral / emotional dimension to the typical bodily, functional, mental and social dimension.

Instruments also vary in their approach. While questionnaire based instruments are more commonly used since they allow the users to answer directly and on their own, others use interview like techniques, which may better comply with the abilities of the user. Using questionnaires in an application (app or web based forms) adds the option of adding or skipping questions due to the answers provided by the user.

In order to assure focus and comparability the HRQoL instruments selected for this paper are of the questionnaire type and address a generic scope or major disease such as cancer [2, 5-7, 12]. The selection follows the usage statistics available in the EU Clinical Trials Register [21] and the reviews of instruments in [18]. Table 2 identifies the selected instruments decided on by both authors for the comparison (rightmost column with the abbreviation used).

**Table 1.** Selected HRQoL instruments

| class type | instrument  | selected as         |
|------------|---|---------------------|
| generic    | Short-Form 36 Health Survey   | SF-36               |
| generic    | Sickness Impact Profile   | -                   |
| generic    | Nottingham Health Profile   | NHP                 |
| generic    | WHOQOL-100 und WHOQOL-BREF  | WHO-100<br>WHO-BREF |
| generic    | EQ-5D-3L und EQ-5D-5L   | EQ-5D-5L            |
| specific   | European Organisation for Research and Treatment of Cancer QLQ-C30                                      | QLQ-C30             |
| specific   | Functional Assessment of Cancer Therapy-General and<br>Functional assessment of Chronic Illness Therapy | FACIT-G             |

To establish relevant criteria for the comparison a literature search was conducted in April 2021 using the following searches without restricting the date of the publication: Pubmed ((Health related Quality of life) OR (Quality of Life)) AND ((instrument)) AND (comparison) with the additional restriction to "systematic review" and Google Scholar "comparison" "instrument" allintitle:("Health related Quality of life" OR "Quality of Life"), resulting in 266 and 248 papers. The successive steps of eliminating the duplicates, using the titles and the abstracts to identify matching papers lead to 36 papers for further reading. Both authors conducted an individual assessment and resolved the ambiguities.

Several papers included comparisons of HRQoL instruments however, these were linked to a specific study and / or disease [15-17, 19-20, 22, 23]. This focus also impacts the identified and applied criteria, which are less overarching as required for the foreseen assessment. However, there are exceptions, for example the publication of Busija [6], which uses a well structured set of criteria to assess HRQoL instruments for rheumatic diseases. Another type of comparison is established by statistics on the usage of instruments, as in [18]. The approaches aiming to establish a framework for assessing HRQoL instruments state several criteria and provide some grouping of them [24, 25]. However, due to their date of publication they do not have clear focus on the ehealth domain. This missing focus becomes also evident when extending the searches above with (ehealth OR e-health) and thereby reducing the number of matching papers to nearly zero.

Based on the analysis of the literature and the specific need for the ehealth domain the following criteria for comparison have been established and structured in six groups (Table 2).

**Table 2.** Criteria derived and the established grouping for assessing HRQoL instruments

| Criteria Group       | Criteria  |
|----------------------|---|
| HRQoL scope          | dimensions covered                                      |
|                      | even / uneven weight of dimensions                      |
|                      | response options (scale, free text)                     |
| practicability       | time required to answer                                 |
|                      | reasonableness  |
|                      | evaluation (manual, automatic)                          |
|                      | management of missing data                              |
| quality              | validity (targeted measurement)                         |
|                      | reliability (changes observed in repeated measurements) |
|                      | objectivity (independence of evaluator)                 |
| ehealth reference    | observation periods (before, during, after)             |
|                      | number of items used to address application acceptance  |
|                      | sensitivity (with and without the ehealth application)  |
|                      | study routine use                                       |
| result comparability | reference data for establishing a baseline              |
|                      | multilingualism / available translations                |
| conditions of use    | license requirement                                     |
|                      | usage fees (academic / commercial use)                  |
|                      | extension modules (e.g. disease specific)               |
|                      | further aspects (e.g. recommended, further parameter)   |

The first group “HRQoL scope” reflects the scope covered by the instrument, in particular the breadth reflected in the dimensions and their relative weights. The group “practicability” includes user and evaluator related criteria referring to the effort needed for providing a response and performing the evaluation. This also includes handling of missing data [17]. The third group addresses “quality” in view of validity, reliability and objectivity [18, 26]. Key items of the “ehealth reference” group are the observation periods and the capability to detect effects, resulting from the ehealth application use, the sensitivity [19]. Others are the number of items and the reported use of the instrument in clinical studies based on an EU wide Clinical Trials Register [21]. For valuing the result of an index based instrument a baseline has to be established which is somewhat collateral with available translations [18], both criteria forming the group “result comparability”. Finally, the last group covers the “conditions of use”, including licenses [26], costs, extensions and recommended use.

For each single criterion a range of 0...2 points has been assigned using documented requirements for achieving 0, 1 or 2 points. For example, the use of a particular instrument in studies: > 500 studies (2 points), 200-499 studies (1 points), less 200 (0 points). In the same line the existence or non-existence of reference data is attributed 2 or 0 points. Initially, the assignment of points had been made by the two authors separately followed by resolving the few differences.

To avoid a bias from the number of criteria per group each group has been separately valued in a range of 0 to 1.0 based on the maximal points achievable per group (achieved points / maximal points). This also allows for a weighing of individual groups using specific scenarios as shown in the following sections.

### 3. Results

Figure 1 exhibits the comparison results for the instruments selected. The maximum score was 6.0. The instruments vary considerably in addressing of HRQoL dimensions and provision of scaled responses. The overall practicability is good. The lower value for WHO 100 results from the large number of items used, which is rated cumbersome for users. For NHP and EQ-5D-5L the evaluation and the handling of missing data is not well established, thus it reduced their score on practicability. Validity as part of the quality group was good for all instruments; however reliability scored less for NHP as well as objectivity for the WHO BREF, WHO 100 and FACIT-G. Clear differences are observed for the reported use in the ehealth domain, with limited use for NHP and WHO 100. The low usage of the instruments NHP, WHO 100 and FACIT-G and limited number of available translations contribute to their low score for the group “result comparability”. Nearly all instruments require a license and expenses which depend on the intended use, be it academic or commercial, however with exception of the WHO 100 and WHO BREF which are free of charge. Those, which scored well in the group “conditions of use”, also provide additional modules for specific aspects.

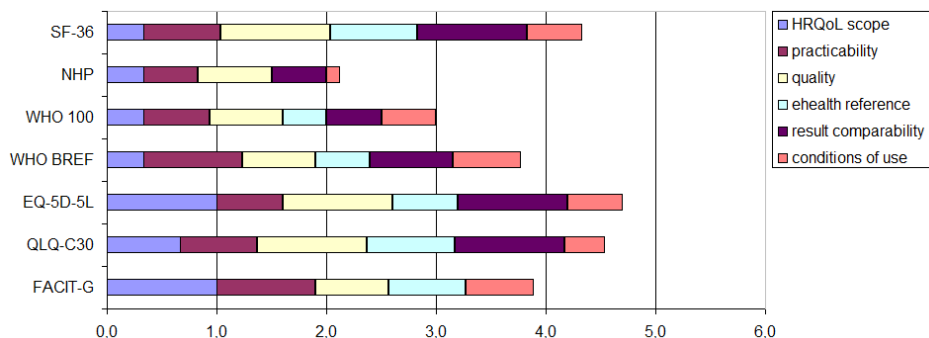


Figure 1. Comparison result for the instruments and the identified groups.

From the generic instruments both the EQ-5D-5L and the SF-36 achieve comparable results, for the specific ones the QLQ-C30 reaches the same level. With a stronger weight on the group ehealth reference e.g. using a factor of 2 this result stays by even reducing the differences in the overall score for each of these instruments.

#### 4. Discussion

The criteria established proved their suitable applicability for assessing HRQoL instruments. When compared to the criteria published by [6] a partial mapping can be established for some groups. However, the illness related group of [6] focuses on strength and weaknesses of the instrument and as such does not have a direct counterpart to a group or criteria identified in this paper. The attributes identified for the “standardized assessment” of the EMPRO tool [25] match well with the criteria of “quality” and “result compatibility” groups and partially with “HRQoL scope” and the “practicability” but the remaining two groups are not covered. The evaluated measurement properties used to derive the COSMIN checklist [26] reveal a different grouping by subsuming criteria from the groups “HRQoL scope”, “quality” and “result compatibility” into the “validity” group identified for COSMIN. Again, there is a limited match for the criteria of the groups “ehealth reference” and the “conditions of use”.

Introducing the group “conditions of use” in this paper is questionable. While from the carers’ point of view the assessment of HRQoL is not directly linked to the licensing and costs of the instrument, the respective organization conducting the review will undoubtedly address these issues. As such this group has a role, which might be valued differently depending on the stakeholder.

The established grouping includes the option of increasing the weight of a specific group to reflect usage requirements. The results (Figure 1) reflect a general scenario with equally weighted groups. When practicability is an issue, e.g. in order to maximize recall based on the ease of use or to assure automated evaluation of the questionnaires, its weight might be increased. In this scenario WHO BREF and FACIT-G would get more attention. In the same line scaling up the impact of the group “ehealth reference” leads to SF-36, EQ-5D-5L and QLQ-C30, to be suitable instruments.

The decision between comparably rated instruments needs to be taken according to the envisaged scope of the HRQoL assessment. For example, both SF-36 and EQ-5D-5L are focused on body function, whereas QLQ-C30 addresses treatment aspects of an illness, in particular cancer. The WHO instruments mainly deal with activities of all day living and personal interaction whereas the FACIT-G values the impact of an illness and user satisfaction.

The instrument NHP has been rated better by other authors [25]; however the limitations observed in most of the groups in this comparison caused the stated result.

None of the instruments was able to score fully in the group “ehealth reference” mainly due to limitations in the observation periods and number of items used to assess the contribution of the ehealth application to HRQoL. This deficit could be remedied by an additional module in particular for those two instruments (QLQ-C30, FACIT-G), which allow for amendments. Such a module could also address application usability, which includes the way the assessment is performed. While a digital version allows omitting irrelevant questions dependent on answers given and allow for easier evaluation, several publications state that paper-based and digital version lead to equivalent results [27].

The comparison of instruments has been based on criteria with guidelines for evaluating. While quantitative properties can be easily evaluated qualitative ones depend to a certain extent on the evaluator. Since this evaluation was only performed by the authors in two steps, starting with an individual assessment and discussion to achieve consensus it may be partially subjective.

## 5. Summary

This paper presents a comparison of HRQoL instruments based on a questionnaire type approach. Criteria have been identified and applied to generic and disease specific instruments which were selected based on their reported usage in studies. The criteria have been structured in six groups (HRQoL scope, practicability, quality, ehealth reference, result comparability, conditions of use), which allow being equally taken into account or for weighted approach to target specific scenarios.

The comparison recommends two generic instruments (SF-36, EQ-5D-5L) and one disease specific instrument (QLQ-C30) giving the best results with little differences. A selection between these three can be easily achieved taking the envisaged scope of the HRQoL assessment into account.

All instruments did not score well with regard to assessing the impact of an ehealth application on HRQoL. Acceptable contribution was achieved for the SF-36, QLQ-C30 and FACIT-G. While the items of the SF-36 are fixed, the other two instruments allow for an amendment by modules, which could be designed to better assess the impact of ehealth applications. The development of such modules could be the topic of further research.

## Contributions of the Authors

Martin Staemmler was in charge of the coordination, review, correction and revision of the manuscript. Marlene Schmidt was in charge of the coordination, study design conception, evaluation and manuscript drafting.

## References

- [1] Geraeds AJLM, Richardson A, Haagsma J, Derret S, Polinder S. A systematic review of studies measuring health-related quality of life of general injury populations update 2010–2018. *Health Qual. Life Outcomes*, vol. 18:160, 2020, pp. 1–20.
- [2] Rautalin M et al. Health-related quality of life in different states of breast cancer – comparing different instruments, *Acta Oncologica*, vol. 57:5, 2018, pp 622-268
- [3] Gallagher AM, Lucas R, Cowie MR. Assessing health-related quality of life in heart failure patients attending an outpatient clinic: a pragmatic approach. *ESC Heart Failure*, vol. 6, 2019, pp 3–9
- [4] Mouelhi Y, Jouve E, Castelli, C Gentile S. How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods, *Health Qual. Life Outcomes*, vol. 18:136, 2020, pp. 1–17.
- [5] Fayers PM, Machin D. Quality of Life: The Assessment, Analysis and Reporting of Patient-reported Outcomes, 3. Edition. Wiley-Blackwell, 2015.
- [6] Busija L, Pausenberger E, Haines TP, Haymes S, Buchbinder R, Osborne RH. Adult measures of general health and Health-Related Quality of Life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQOL). *Arthritis Care Res.*, vol. 63, no. S11, Nov. 2011, pp. 383–412.
- [7] Chiarotto A, Terwee CB, Kamper SJ, Boers M, Ostelo RW. Evidence on the measurement properties of health-related quality of life instruments is largely missing in patients with low back pain: A systematic review. *J. Clin. Epi.*, vol 102, 2018, 23-37
- [8] RAND Cooperation, 36-Item Short Form Survey (SF-36), Available from [https://www.rand.org/health-care/surveys\\_tools/mos/36-item-short-form.html](https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html) (accessed 5 April 2021).

- [9] Hunt SM, McKenna SP, McEwen J, Williams J, Papp E. The Nottingham health profile: Subjective health status and medical consultations. *Soc. Sci. Med. Part Med. Psychol. Med. Sociol.*, vol. 15, no. 3, Part 1, May 1981, pp. 221–229.
- [10] World Health Organization. WHOQOL: Measuring Quality of Life. Available from <https://www.who.int/tools/whoqol> (accessed 5 April 2021).
- [11] van Hout B et al. Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets. *Value Health*, vol. 15, no. 5, Jul. 2012, pp. 708–715.
- [12] Fayers P, Bottomley A. Quality of life research within the EORTC—the EORTC QLQ-C30. *Eur. J. Cancer*, vol. 38, Mar. 2002, pp. 125–133.
- [13] Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. *Health Qual. Life Outcomes*, vol. 1, no. 79, Dec. 2003.
- [14] Abma IL, Butje BJD, Ten Klooster PM, van der Wees PJ. Measurement properties of the Dutch-Flemish patient-reported outcomes measurement information system (PROMIS) physical function item bank and instruments: a systematic review. *Health Qual Life Outcomes*. vol. 19:62, 2021, pp 1–22
- [15] Estoque RC et al. A review of quality of life (QOL) assessments and indicators: Towards a ‘QOL-Climate’ assessment framework, *Ambio*, vol. 48, no. 6, Jun. 2019, pp. 619–638.
- [16] Pequeno NPF et al. Quality of life assessment instruments for adults: a systematic review of population-based studies. *Health Qual. Life Outcomes*, vol. 18, no. 208, 2020, pp. 1–13.
- [17] Bourdel N et al. Systematic review of quality of life measures in patients with endometriosis. *PLOS ONE*, vol. 14, no. 1, Jan. 2019, pp. 1–32.
- [18] Haraldstad K et al. A systematic review of quality of life research in medicine and health sciences. *Qual. Life Res.*, vol. 28, no. 10, Jun. 2019, pp. 2641–2650.
- [19] Marra CA et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Soc. Sci. Med.*, vol. 60, no. 7, Apr. 2005, pp. 1571–1582.
- [20] O’Carroll RE, Smith K, Couston M, Cossar JA, Hayes PC. A comparison of the WHOQOL-100 and the WHOQOL-BREF in detecting change in quality of life following liver transplantation. *Qual. Life Res.*, vol. 9, Feb. 2000, pp. 121–124.
- [21] EU Clinical Trials Register. Available from <https://www.clinicaltrialsregister.eu> (accessed 5 April 2021)
- [22] Wohlleber K, Heger P, Probst P, Engel C, Diener MK, Mihaljevic AL. Health-related quality of life in primary hepatic cancer: a systematic review assessing the methodological properties of instruments and a meta-analysis comparing treatment strategies. *Qual Life Res.*, 2021, online ahead of print
- [23] Jay CL, Butt Z, Ladner DP, Skaro AI, Abecassis MM. A review of quality of life instruments used in liver transplantation. *J. Hepatol.* Vol 51 no 5, 2009, pp 949-59
- [24] Mokkink LB, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J. Clin. Epi.*, vol 63, 2010, 737-745
- [25] Valderas JM, et al. Development of EMPRO:A Tool for the Standardized Assessment of Patient-Reported Outcome Measures, *ISPOR*, vol. 11, no 4, 2008, pp 700-708
- [26] Chen T, Li L, Kochen MM. A systematic review: How to choose appropriate health-related quality of life (HRQOL) measures in routine general practice? *J. Zhejiang Univ. Sci.*, vol. 6B, no. 9, 2005, pp. 936–940.
- [27] Muehlhausen W et al. Equivalence of electronic and paper administration of patient-reported outcome measures: A systematic review and meta-analysis of studies conducted between 2007 and 2013. *Health and Quality of Life Outcomes*, 13(167), 2015.