Abstract

HR-QoL measurements attempt to turn subjective information into objective information. In this article, I describe the different kinds of health-related quality of life (HR-QoL) measures, how they work, and how they can be interpreted. Main types of HR-QoL measures include generic; disease- or population-specific; dimension-specific; individualised; and preference-based. Each serves different purpose and should be applied to different populations. For example, generic measures can be used in general populations and across various diseases, whereas disease-specific ones address specific diseased populations. I also discuss key considerations for using and presenting HR-QoL measures, including ensuring that a validated and legally obtained measure is administered; describing the type and specifics of the HR-QoL measure; and explaining how the measure was used, how scores were computed, and how to interpret them.

Generic HR-QoL measures

Generic HR-QoL measures are designed for use in any population, irrespective of disease status, that is, in patients regardless of the condition they suffer from and in general populations. Many generic measures focus on physical function and measure impairment, disability, or handicap. Others cover psychological issues. Although they are often considered as not being sensitive enough to detect changes specific to certain diseases, they allow comparisons across different conditions and with general populations. The most widely used generic measures are EQ-5D, the SF-36 (36-Item Short Form Survey) and the NHP (Nottingham Health Profile).

EQ-5D

The EQ-5D defines five dimensions of health: mobility, self-care, usual activities, pain/discomfort, and anxiety or depression. In the original version of the EQ-5D, currently referred to as EQ-5D-3L, each dimension is categorised into three levels of burden: 1) no problem, 2) a moderate problem, and 3) an extreme problem. Furthermore, some measures can fall into two categories; for example, dimension-specific measures, which can be used in general population as well as across different diseases, can also be considered generic.
and is being used in many clinical and economic studies as well as population surveys all over the world. Over the years, the measure has evolved and now two other versions are available – the EQ-SD-5L (Figure 1) and the EQ-SD-Y. The EQ-SD-5L was introduced in 2009 to increase sensitivity and reduce ceiling effect over the EQ-SD-3L. It contains two intermediate categories of burden: slight and severe. The EQ-SD-Y targets children and adolescents aged 8 to 15 years and is also available as a proxy measure. All EQ-SD measures can be administered as a paper or electronic version. To use any of EQ-SD measures, a planned study or project needs to be registered at https://euroqol.org/support/how-to-obtain-eq-5d/ and the conditions of use agreed upon with the EuroQol EuroQol Research Foundation Office.

**SF-36**

The SF-36 measures physical and mental health as well as provides assessment of general health. Physical health includes physical functioning (10 items), physical role functioning (4 items), bodily pain (2 items), and general health (6 items). Mental health includes vitality (4 items), emotional role functioning (3 items), social role functioning (2 items), and mental health (5 items). For most items, Likert’s scale is used. The SF-36 is available in shorter versions, such as SF-6D, SF-12, and SF-20, of which the SF-6D is used primarily in health economic evaluations. These measures are in the public domain and free-of-charge, although certain legal conditions are imposed, for example, proper acknowledgement. They can be downloaded from https://www.rand.org/health/surveys_tools.html. Following the instructions for calculating the scores is crucial because items for physical and mental health are constructed in opposite directions. The raw scores from the SF-36 can be standardised on a 100-point scale, assuming equal weighting for each item. For some countries, such as Germany, country-specific weights are available and should be used for national data. Overall, a lower score denotes poorer HR-QoL.

**Nottingham Health Profile**

The NHP (Figure 2) is another example of a generic measure. It focuses on feelings and emotions, rather than physical performance, and is includes 38 items (statements) in six dimensions, as explained in the accompanying article “Measuring Quality of Life – theoretical background” in this issue of Medical Writing (page 8). The respondent selects “yes” or “no” according to whether a certain problem applies. The score is calculated by adding the number of “yes” answers (i.e., the number of recognised problems). Thus, a higher score denotes poorer HR-QoL. Galen Research is the copyright holder and should be contacted at http://www.galen-research.com to request permission for its use.

**Disease/population-specific measures**

Disease-specific measures are developed to address the need to monitor patients with increased accuracy and to provide enough sensitivity to detect features of specific conditions. Currently, many disease-specific measures targeting various patient populations are available.

**EORTC QLQ-C30**

One of the first disease-specific HR-QoL measures is the EORTC QLQ-C30. This measure targets various patient populations and assesses HR-QoL for cancer patients. It consists of 30 items across nine domains: Swallowing difficulties, nausea and vomiting, tiredness, pain, sleep problems, cognitive function, emotional function, social function, and global health. Each item is rated on a five-point scale ranging from “not at all” to “very much”.

**Figure 1. EQ-SD-5L – sample**


**Figure 2. Nottingham Health Profile (NHP) – sample**

Quality of life measures – an overview – Kołtowska-Häggström

Figure 3. European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 – sample
Reprinted with permission from the European Organisation for Research and Treatment of Cancer. For permission to use, contact the Quality of Life Department at EORTC (http://qol.eortc.org/)

C30 is modular: the core addresses the issues generally encountered by patients with any cancer, and specific modules are included for different types of cancer or their treatment.

AcroQoL
Another example of a disease-specific measure is the AcroQoL (Acromegaly Quality of Life Questionnaire). It contains 22 items describing problems experienced by patients with acromegaly. The items cover three dimensions: physical, psychological/appearance, and psychological/personal relations. Responses are based on Likert’s scale, and depending on the item, choices for frequency or the degree of agreement with the problem described, are coded from 1 to 5. The raw scores for each item are summed and then standardised to a scale of 22 to 110, with a higher score corresponding to a better HR-QoL.

HR-QoL in children and adolescents: KINDL and CAT-SCREEN
Interest in HR-QoL measures for the paediatric population is growing, and many measures have been or are being developed. KINDL®, originally developed by Monika Bullinger in 1994 assesses HR-QoL in healthy or diseased children and adolescents aged 3 to 17 years. In addition to the core generic module, several disease-specific modules have been developed, such as for paediatric patients with asthma, epilepsy, cancer, diabetes, or obesity. The core measure contains 24 items and is provided in three versions for different age groups (4-6, 7-13, and 14-17 years), each of which can be completed by a child or adolescent and their caregiver. KINDL is available as a paper-pencil version and an electronic version called CAT-SCREEN (Figure 4). All versions are copyrighted.

Dimension-specific measures
Dimension-specific measures focus on certain HR-QoL domains, such as pain, fatigue, and anxiety and depression. Examples include HADS (Hospital Anxiety and Depression Scale), the McGill Pain Questionnaire, and the MFI (Multidimensional Fatigue Inventory). The structure and principles of dimension-specific measures are similar to those of disease-specific ones, as described above. Depending on the nature of the items, these measures can be used for any diseased population or even for a general population and thus could be considered generic.

Individualised measures
Individualised measures aim at evaluating HR-QoL from respondents’ own perspective and allow them to either include items of their choice...
Figure 5. QLS-H, an example of an individualised HR-QoL measure


and allocate weights or to only allocate weights for predefined items. In the case where respondents include items of their own choice, they first select the most important issues relating to their HR-QoL. (step 1) and then self-rate the level of problems they face (step 2). After this, they allocate weights to them (step 3). In the case where respondents use predefined items, only steps 2 and 3 are followed. The SEIQoL (Schedule for the Evaluation of Individual Quality of Life)\(^{19}\) and PGI (Patient Generated Index),\(^{20}\) which use all three steps, laid the groundwork for individualised measures. The administration manual for the SEIQoL, published in 1993 by O’Boyle and colleagues, describes the whole process in detail.\(^{21}\) In principle, the scores for each item are calculated by multiplying self-ratings by allocated weights. The sum of calculated scores for each item comprises the final score (index). The QLS-H (Questions on Life Satisfaction Modules-Hypopituitarism), developed for adult patients with growth hormone deficiency, is an example of a disease-specific, individualised measure containing predefined items (Figure 5).\(^{22}\)

Preference-based (utility) measures

Preference-based measures emerged from decision-making theory and are mainly used in pharmacoeconomic evaluations, also known as cost-utility analyses. The basic requirement is to incorporate patient or general population weights (utilities) for different health states assigned under uncertainty.\(^{23}\) Utilities range from 0 (death) to 1 (perfect health), although negative numbers are possible for states considered worse than death. Utilities are used to derive QALY (quality-adjusted-life-years). A number of techniques are used to generate utilities,\(^{24}\) such as time trade-off, as used in EQ-SD, referred to as the EQ-VT approach;\(^{25}\) standard gamble, as used in the SF-6D (Short Form 6D);\(^{26}\) or VAS with relevant anchors. Briefly, Time trade-off asks respondents to decide how many years of life in a described (given) condition they are prepared to give up in order to live in full health. In other words, they are asked if they prefer to live shorter in full health instead of living a certain number of years longer in a given health state or condition. Standard gamble presents alternative treatments with probabilities of better and poorer outcome to life in given health state or condition. Responders provide the highest acceptable risk of treatment failure (e.g. death). Standard gamble and time trade-off are the gold standards for measuring health utilities, but they can also be generated using a combination of standard gamble and multi-attribute scaling analysis,\(^{24}\) as in the HUI2 and HUI 3 (Health Utilities Index 2 and 3), or based on the SF-6D and the EQ-5D.

Conclusion

HR-QoL is an important construct widely used in daily patient management, clinical trials, health economics and medical decision making. Each of these applications imposes different requirements on the HR-QoL measures. Clinical use usually requires a measure that captures specific changes within a certain disease, within a patient population (in clinical trials), or for individual patients (in daily clinical practice). Pharmacoeconomic evaluation often requires that health status is expressed as a single summary score (a health status index) capable of identifying and quantifying differences across diseases and aggregate changes in patient health status over time. This explains why so many HR-QoL measures have been developed.

When working with HR-QoL data and writing manuscripts or other documents, medical writers should keep in mind the following:

1. Most scales used in HR-QoL measures are ordinal, meaning that categories are not equally spaced. For example, the distance from “not important at all” (1) to “little important” (2) is not necessarily the same as between “little important” (2) to “important” (3). This explains why so many HR-QoL measures have been developed.

2. Responses and thus scores are subjective, meaning that the values behind them differ from one another. This depends on many different factors, such as personality, health and overall life experiences, and cultural norms.

3. Understanding how a measure is constructed and how answers (choices) are coded is important when writing about them. For example, is a higher numerical score better or worse, and does an increase in score indicate improvement or deterioration?
When writing, be sure to explain how to interpret the scores.

4. When comparing results originating from different HR-QoL measures, check whether they are based on working scores or scale scores.

5. Make sure that the researchers used a legal version of a measure and that proper acknowledgement is included. If a measure is publicly available (i.e. no licence needed), be sure to state so and acknowledge the source of the measure. Also, include information about the version number and the mode of administration needs.

6. In cases where a translation of a measure is used, confirm that it was properly translated and validated, and provide a few lines about it in the manuscript.

7. For manuscripts, follow the 2013 CONSORT-PRO extension while presenting data from clinical trials that include patient-reported outcome measures.

Acknowledgements

I would like to thank Claire Gudex and Diarmuid De Facoite for their helpful comments on the first draft of this article. I also would like to thank Phil Leventhal for all his linguistic corrections and edits in my article.

Conflicts of interest

The author declares no conflicts of interest.

References


20. Ruta DA, Garratt AM, Leng M, Russell IT, MacDonald LM. A new approach to the measurement of quality of life. The Patient-Generated Index. Med Care. 1994;32(11):1109–26.


Clinical use usually requires a measure that captures specific changes within a certain disease, within a patient population (in clinical trials), or for individual patients (in daily clinical practice).


Author information

Maria Kołtowska-Häggström, MD, PhD, a co-owner of Proper Medical Writing, the first Polish medical writing agency, earned her PhD at Uppsala University (2008) based on the dissertation "Quality of life in adult patients with growth hormone deficiency; bridging the gap between clinical evaluation and health economic assessment". Maria is author of over 70 peer-reviewed publications, many of them relating to research on patient-reported outcomes. She is also a member of the European Association of Scientific Editors, European Society of Endocrinology, and Growth Hormone Research Society; reviewer for a number of journals; and an Associate Editor for BMC Endocrine Disorders. A longtime EMWA member, Maria is a section editor for Medical Writing, workshop leader, and Chair of an Expert Seminar Series.