

Refereed article

Comparison of *howRU* and EQ-5D measures of health-related quality of life in an outpatient clinic

Tim Benson

Routine Health Outcomes Ltd., 14 Pinewood Crescent, Hermitage, Thatcham RG18 9WL, UK
UCL Centre for Health Informatics & Multiprofessional Education (CHIME), Wolfson House, 4 Stephenson Way, London, NW1 2HE, UK

Henry W W Potts

UCL Centre for Health Informatics & Multiprofessional Education (CHIME), Wolfson House, 4 Stephenson Way, London, NW1 2HE, UK

Justin M Whatling

UCL Centre for Health Informatics & Multiprofessional Education (CHIME), Wolfson House, 4 Stephenson Way, London, NW1 2HE, UK

David Patterson

The Whittington Hospital, Magdala Avenue, London N19 5NF, UK

Cite this article: Benson T., Potts HWW., Whatling JM., Patterson D. Comparison of *howRU* and EQ-5D measures of health-related quality of life in an outpatient clinic. *Inform Prim Care*. 2013;21(1):12–17.

<http://dx.doi.org/10.14236/jhi.v21i1.9>

Copyright © 2013 The Author(s). Published by BCS, The Chartered Institute for IT under Creative Commons license <http://creativecommons.org/licenses/by/2.5/>.

Author address for correspondence:

Tim Benson
Routine Health Outcomes Ltd.
14 Pinewood Crescent, Hermitage,
Berkshire RG18 9WL, UK
Email: tim.benson@howrwe.com

Accepted October 2013

ABSTRACT

This paper reports on a head-to-head study of *howRU* and EQ-5D on patients with cardiovascular disease. *howRU* is a short generic measure of health-related quality of life comprising 39 words, designed for routine use, which we compare with EQ-5D (230 words). Patients attending a clinic completed both instruments. Completed data were available for 116 patients, 51% female, mean age 56 and SD 20. *howRU* is shorter, has better readability statistics, a higher completion rate, a wider range of states used and a smaller ceiling effect than EQ-5D. The correlations of *howRU* with EQ-5D are similar to those of EQ-5D with other validated instruments.

Keywords: EQ-5D, health-related quality of life, health status, *howRU*, patient-reported outcomes, PROMs

BACKGROUND

Measures of health-related quality of life (HRQoL), also known as patient-reported outcome measures (PROMs), were developed originally for population health surveys, economic evaluation of new medicines and resource allocation. More recently, they have been used to monitor the performance of individual clinical teams.¹ PROMs form an overarching indicator in the National Health Service (NHS) outcome framework,² reinforcing the NHS' intention to move away from focusing on activity and process targets to measuring outcomes as perceived by patients.

Practicality is a limiting factor on the routine adoption of PROMs at the point of care.³ In particular, well-established

measures place a significant burden on respondents in terms of the time taken to complete forms and their understandability.

This paper describes a short study in which two of the shortest generic instruments, *howRU* and EQ-5D, were completed in a clinical setting by the same cohort of patients. We compare the questionnaire length, readability, completion rate, distribution of states used including ceiling and floor effects, and the level of agreement.

This study is of interest because EQ-5D is widely used in the NHS, while *howRU* is a newer instrument that is designed to be short, quick and easy to use.

howRU

howRU has been developed to meet the need for a quick, easy-to-use, generic PROM for frequent repeated use.⁴ The objective was to provide standardised, comparable measure of health outcomes that are clinically relevant, sensitive, responsive to change and easy to interpret, applicable to all patients with any combination of conditions, across all care sectors, including self-care, primary, secondary, community and social care. The design criteria include brevity, understandability, simplicity and suitability for electronic data collection using mobile phones and tablets.

The design was constrained by the size of a smartphone screen. Figure 1 shows the form used in the survey. The *howRU* questionnaire has a top-level question *How are your today?* and four items:

- **Pain or discomfort**—physical symptoms.
- **Feeling low or worried**—emotional symptoms.
- **Limited in what I can do**—work, home and leisure activities.
- **Dependent on others**—need for help from others (since the data reported here were collected, this has been modified to *Require help from others*).

Each item has four options, with labels: **none**, **slight** (since the data reported here were collected, this has been modified to *a little*), **quite a lot** and **extreme**. The labels are listed in order of severity, are colour-coded like traffic lights (green, yellow, orange and red) and annotated with smiley pictographs. The labels, order, colours and pictographs indicate severity in complementary ways. The resulting matrix provides 256 health states in all (4⁴).

The *howRU* score is calculated by adding the values for each level, using *none* = 3, *slight* = 2, *quite a lot* = 1 and *extreme* = 0, giving a 13-point scale with a range from 0 (worst) to 12 (best). Sub-scales may also be provided for discomfort, distress, disability and dependence. For population studies, the mean scores are transformed to a scale from 0 (all *extreme*) to 100 (all *none*).

Circle one face on each line to tell us how you are today

How are you today?	None	Slight	Quite a lot	Extreme
Pain or discomfort				
Feeling low or worried				
Limited in what I can do				
Dependent on others				

© 2008,2009 Routine Health Outcomes Ltd. All rights reserved. 11/09 howRu™

Figure 1 *howRU* form

EQ-5D

EQ-5D is a standardised generic self-completion measure of health status, originally developed for clinical and economic

appraisal and population surveys, and in particular to support the calculation of quality-adjusted life years (QALYs) in cost-utility analysis.⁵ EQ-5D has been used in the NHS PROMs programme for hip and knee replacement, groin hernia and varicose vein operations.

EQ-5D has two main parts: the descriptive system together with its associated weighting system and a visual analogue scale (EQ-VAS).

The EQ-5D descriptive system has five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each with three response options, typically (1) none/absent, (2) moderate/some and (3) unable/extreme, creating 243 (3⁵) possible states.

For example, a patient with migraine might report a score such as 11231 meaning:

- 1—I have no problems in walking about.
- 1—I have no problems with self-care.
- 2—I have some problems with performing my usual activities.
- 3—I have extreme pain or discomfort.
- 1—I am not anxious or depressed.

A value set (also known as a tariff) is used to weight each possible health state in the descriptive system, using a scale from 1 (best conceivable health state) through dead (value 0) to the worst conceivable health state (which may have a negative value).⁶ The resulting score is the EQ-5D Index.

This study used the preference-based tariff for the UK general population.⁷ Weights are subtracted from a maximum value of 1.0 using the values shown in Table 1.

Table 1 EQ-5D weights

Item score	Mobility	Self-care	Usual activity	Pain/discomfort	Anxiety/depression
2 (some problem)	-0.069	-0.104	-0.036	-0.123	-0.071
3 (extreme problem)	-0.314	-0.214	-0.094	-0.386	-0.236

If one or more dimensions indicates either a moderate problem (score 2) or an extreme problem (score 3) then an additional constant of -0.081 is subtracted; if one or more dimensions is scored 3, a further constant of -0.269 is subtracted.

The value for profile used in the example above (11231) is therefore:

$$1.0 - 0.081 (\text{score of } 2) - 0.269 (\text{score of } 3) - 0.036 (\text{usual activity } 2) - 0.386 (\text{pain/discomfort } 3) = 0.228.$$

When used in QALY calculations, this is interpreted as meaning that 4.4 years in this state are considered to be equivalent to one year in good health (1 divided by 0.228 years).

The second part of the instrument, the EQ-VAS, is a 20 cm vertical thermometer anchored at 0 (worst imaginable health state) and 100 (best imaginable health state). The patient marks their present state by drawing a line from a box marked *Your own health today* to the VAS.

METHOD

This study was undertaken in the cardiovascular clinic at the Whittington Hospital in North London, during July and August 2009. The hospital's catchment area is a socially mixed, ethnically diverse and inner city area. Many patients were born outside of UK.

Ethics approval was obtained from the UCL/UCLH Research Ethics Committee (08/H0715/94). Participation was voluntary.

The number of words in each questionnaire was counted, including instructions.

The readability of both instruments (including instructions for use) was estimated using the Flesch–Kincaid grade level and the Flesch Reading Ease score, included as a part of Microsoft Word for Mac 2008.

A consecutive series of patients were approached and asked to participate in an evaluation of two instruments that might be used to measure quality of life for patients. Each willing participant completed a four-page booklet. The order of the instruments in the booklet was alternated for different subjects to avoid and allow testing for order bias.

Each subject recorded their age, sex and whether their first language was English. The researcher checked each returned form for completeness; if any missing data were detected, and the subject was still present in the waiting area, then he or she was asked to complete it; this reduced the missing data rate. No record was kept of what items were missed initially. In practice, the main reason for incomplete entries was that patients had left the waiting area, because they had been called in to see a clinician.

The completion rates and frequency distributions of the two instruments were compared.

To test construct validity, we predicted that:

- The severity of physical aspects would correlate within and between the instruments, and that emotional aspects would form a separate grouping.
- Both instruments would find that older subjects would have more severe states but neither instrument would show differences between subjects according to sex, first language or order of presentation.
- We would find high correlations (in the order of $r = 0.70$ or above) between summary scores for each instrument: the aggregate *howRU* score, the EQ-5D Index and the EQ-VAS.

RESULTS

Readability and length

EQ-5D is longer (230 words) including instructions than *howRU* (39 words).

Readability statistics (Flesch–Kincaid Grade and Flesch Reading Ease scores) are shown in Table 2. The Flesch–Kincaid grade score estimates the number of years of schooling required and the Flesch Reading Ease score provides a reading ease score between 0 and 100 (higher is better).⁸ *howRU* has a lower reading age (better readability).

Table 2 Length and readability data

	Total number of words including instructions	Flesch–Kincaid grade	Flesch reading ease
<i>howRU</i>	39	2.6	86
EQ-5D	230	6.0	71

Completion

A total of 150 patients were approached and 125 participants agreed to participate. Almost all of those who declined to participate spoke a little or no English.

Five subjects (4%) completed *howRU* but did not complete all of EQ-5D. All who completed EQ-5D also completed *howRU*. This difference is statistically significant (Fisher exact test, $p < 0.001$).

Less than half of the subjects (56/116, 48%) followed the instructions for completing the EQ-VAS, as stated on the form. (These instructions ask the subject to draw a line from the box marked *Your own health today* to the VAS.) However, the intended answer was clear in every case and these subjects are included in the analysis.

Four subjects (3%) completed the cover sheet only. Subjects with any missing data were excluded from further analysis, leaving 116 subjects with complete data.

Frequency distributions

Completed forms came from 57 men (49%) and 59 women (51%). The mean age was 56 years (men 60 and women 53), standard deviation 20 years and range 15–95 years.

The frequency distribution for *howRU* ratings is shown in Table 3. The highest level of problems reported was for *Limited in what I can do* (59%); the lowest level was *Dependent on others* (34%); 43 out of the 256 possible *howRU* states (17%) were used and 25 subjects (22%) reported no problems on any dimension (the ceiling state). No subjects were in the floor state (all *Extreme*).

Table 3 *howRU* frequency distribution ($N = 116$)

	None	Slight	Quite a lot	Extreme
Pain or discomfort	55 (48%)	48 (41%)	12 (10%)	1 (1%)
Feeling low or worried	52 (45%)	50 (43%)	11 (10%)	3 (3%)
Limited in what I can do	47 (41%)	47 (41%)	19 (16%)	3 (3%)
Dependent on others	76 (66%)	25 (22%)	11 (10%)	4 (3%)

The frequency distribution for EQ-5D ratings is shown in Table 4.

Table 4 EQ-5D frequency distribution (N = 116)

	N (%)
<i>Mobility</i>	
I have no problems in walking about	68 (59%)
I have some problems in walking about	48 (41%)
I am confined to bed	0 (0%)
<i>Self-care</i>	
I have no problems with self-care	101 (87%)
I have some problems washing or dressing myself	13 (11%)
I am unable to wash or dress myself	2 (2%)
<i>Usual activities</i>	
I have no problems with performing my usual activities	60 (52%)
I have some problems with performing my usual activities	45 (39%)
I am unable to perform my usual activities	11 (10%)
<i>Pain/discomfort</i>	
I have no pain or discomfort	54 (47%)
I have moderate pain or discomfort	57 (49%)
I have extreme pain or discomfort	5 (4%)
<i>Anxiety/depression</i>	
I am not anxious or depressed	67 (58%)
I am moderately anxious or depressed	45 (39%)
I am extremely anxious or depressed	4 (3%)

The highest level of problems reported in EQ-5D was for *Pain/discomfort* (53%); the lowest level was *Self-care* (13%); 31 out of the 243 possible states (13%) were used and 27% of subjects reported no problems on any dimension (11111, the ceiling). No subjects were in the floor state (33333).

Eighteen subjects (16%) reported no problems on any item in both *howRU* and EQ-5D, the ceiling state on each instrument. Thirteen (11%) scored at the ceiling in EQ-5D but not *howRU*. Seven (6%) scored at the ceiling in *howRU* but not EQ-5D. This difference, that *howRU* has a smaller ceiling effect than EQ-5D, is significant (Fisher exact test, $p < 0.001$).

Age in years is moderately negatively correlated (Spearman rank-order correlation) with the *howRU* score ($r_s = -0.41$, $p < 0.001$), the EQ-5D Index ($r_s = -0.39$, $p < 0.001$) and the EQ-VAS ($r_s = -0.38$, $p < 0.001$), showing that patient health status reduces with age.

Sex, English first language and order of presentation of instruments are not significantly related to any variables (Mann–Whitney *U* tests).

The relative magnitudes of Spearman correlations between EQ-5D and *howRU* items are all as expected (Table 5). EQ-5D mobility correlates mostly with *howRU limited in what you can do* ($r_s = 0.62$); EQ-5D *self-care* correlates mostly with *howRU dependent on others* ($r_s = 0.51$); EQ-5D *usual activities* correlates mostly with *howRU limited in what you can do* ($r_s = 0.64$); EQ-5D *pain/discomfort* correlates mostly with *howRU pain or discomfort* ($r_s = 0.60$) and EQ-5D *anxiety/depression* correlates mostly with *howRU feeling low or worried* ($r_s = 0.40$).

Table 5 Spearman correlations between EQ-5D dimensions and *howRU* items

<i>howRU</i> item EQ-5D dimension	Pain or discomfort (r_s)	Feeling low or worried (r_s)	Limited in what you can do (r_s)	Dependent on others (r_s)
Mobility	0.48	0.15	0.62	0.51
Self-care	0.27	0.11	0.42	0.51
Usual activities	0.57	0.17	0.64	0.55
Pain/discomfort	0.60	0.26	0.45	0.44
Anxiety/depression	0.28	0.40	0.09	0.16

For individual subjects, the Pearson correlation matrix of the *howRU* score, EQ-5D Index and EQ-VAS is shown in Table 6. The EQ-VAS correlations were lower than originally anticipated. All correlations are significant $p < 0.0001$.

Table 6 Correlation matrix for *howRU*, EQ-5D Index and EQ VAS

	<i>howRU</i>	EQ-5D Index	EQ VAS
<i>howRU</i>	1.00	0.70	0.60
EQ-5D Index	0.70	1.00	0.50
EQ VAS	0.60	0.50	1.00

Table 7 shows the EQ-5D Index mean and standard deviation for all responses with the same *howRU* score. With the exception of the singleton with *howRU* score of 3, the mean values are well correlated.

Table 7 EQ-5D mean and standard deviation for each *howRU* score

<i>howRU</i> score	n	EQ-5D Index mean	EQ-5D standard deviation
12	25	0.94	0.11
11	17	0.85	0.13
10	17	0.81	0.14
9	23	0.73	0.14
8	11	0.65	0.28
7	7	0.38	0.27
6	6	0.46	0.21
5	7	0.40	0.26
4	2	0.27	0.34
3	1	0.69	–
Total	116	0.73	0.25

The standard deviations of EQ-5D, shown in Table 7, fall into two groups: in the range of *howRU* scores from 9 to 12, the EQ-5D standard deviations are in the range 0.11–0.14, but in the more severe range, with *howRU* scores < 9, the standard deviations are much higher, in the range 0.21–0.34. This is a consequence of the EQ-5D tariff, which has a large

decrement if any dimension is rated as severe (see Table 1 and associated description above).

DISCUSSION

A summary comparison of *howRU* and EQ-5D is shown in Table 8.

Table 8 Comparison of *howRU* and EQ-5D properties

Aspect	<i>howRU</i>	EQ-5D
Original design purpose	Routine outcomes monitoring	Economic evaluation
No. of items	4	5 plus VAS
Options for each item	4	3
Scoring scheme	Summative score	Utility weightings
No. of words (including instructions)	39	230
Readability (grade/score)	2.6/86	6.0/71

howRU is shorter (39 words) than EQ-5D (230 words). McDowell⁹ described EQ-5D as the *shortest and simplest of the leading generic instruments for measuring HRQoL of life*. The brevity of *howRU* is important because questionnaire length determines how long it takes for each patient to read and complete the questionnaire.

Readability statistics are a proxy for understandability. This is also important because patients are concentrated amongst the most vulnerable, disadvantaged and deprived sectors of society. Paz *et al.*¹⁰ suggested a criterion of five years of schooling as appropriate for health surveys, and found that none of seven commonly used surveys that they tested, including EQ-5D, achieved this. The readability grade, which approximates to years of schooling needed, is 2.6 for *howRU*, compared with 6.0 for EQ-5D.

The original design objectives for *howRU* and EQ-5D were different. *howRU* was designed for routine and repeated use across all care sectors as part of day-to-day patient care and self-care. For this purpose, usability criteria, including understandability, speed and ease of use are paramount. EQ-5D was originally designed for economic evaluation, for use in clinical trials and population surveys as a way to estimate QALYs. It has been used in routine care mainly because nothing shorter was available.

howRU had a significantly higher completion rate than EQ-5D ($p < 0.01$). It also used a wider range of states and has a smaller ceiling effect than EQ-5D.

The correlations for individual patients between the *howRU* score and EQ-5D Index ($r = 0.70$) are similar to those found in a study, which compared EQ-5D with four other instruments (SF6D, AQoL, HUI3 and 15D).¹¹

The correlation between the EQ-5D Index and the EQ-VAS was unexpectedly low ($r = 0.50$). However, this is consistent with the correlation of $r = 0.45$ between EQ-5D Index and EQ-VAS in pre-operative patients in 294,000 patients taking part in the NHS PROMs programme.^{12: Table 3}

We have considered reasons for the relatively low correlations with EQ-VAS. The EQ-VAS asks subjects to rate their health in general on a 0–100 scale. This is likely to be impacted by their expectation of prognosis. On the other hand, both EQ-5D (five dimensions) and *howRU* ask subjects about their health status today, which is a different question. Furthermore, in a clinic, some of the subjects may feel relatively well, but have a serious underlying illness and might score relatively low on EQ-VAS; others may feel poorly yet expect to make a complete recovery and so score more highly on EQ-VAS.

In our study, only 48% of respondents completed the EQ-VAS in the way instructed. This is similar to the comparable proportion (45%) found by Feng *et al.*,¹² who have explored the performance of the EQ-VAS in greater depth.

The most obvious limitations of this study derive from its modest sample size ($n = 116$), but this is sufficiently large for the statistical tests used. The location at a cardiovascular clinic may mean that results obtained here would not be the same in other populations, although our primary purpose was to compare the instruments head-to-head. There is the further limitation that this study was performed by the originator of *howRU*, but care was taken to minimise risks of bias from this source.

The version of the questionnaire used in this study (data were collected during July and August 2009) differs from the current 2013 version in three small details; the label *slight* has been changed to *a little*, which can to time as well as quality; the phrase (*24 hours*) has been added after the main question *How are you today?* to make it clear that today applies to a full 24 hours, not just right now and the phrase *Dependent on others* has been clarified to *Require help from others*. These differences are not considered to be significant or to have affected the findings.

Since the data for this study were collected, the Euroqol Group has published a revised version of EQ-5D, known as EQ-5D-5L, with five possible responses for each item and other changes.¹³ EQ-5D-5L is longer than EQ-5D (3L) (286 words vs 230) and readability is better (Flesch-Kincaid Grade 3.1 vs 6.0). New methods are also being developed to address some of the limitations of the existing values sets.¹⁴

On each of the relevant comparisons analysed in this paper, the performance of *howRU* has been found to show better properties for repeated routine use than EQ-5D. There is a need for more research on *howRU*, including test–retest, before and after and longitudinal studies across a range of conditions, care settings and demographic groups.

In this and in previous studies, it has been shown that *howRU* performs at least as well as longer generic measures, yet it is considerably shorter and simpler. Where the requirement is to measure patients' health-related quality of life, and selection criteria include universality, ease of use and brevity, then *howRU* may be an appropriate choice.

Contributions

All authors designed the study and contributed to the final text. Tim Benson collected the data and wrote the paper. Tim Benson developed *howRU* with help from Justin M Whatling. Henry W W Potts performed the initial statistical analysis and obtained ethics approval. David Patterson facilitated the data collection.

Competing interests

Tim Benson is the founder of Routine Health Outcomes Ltd., which owns the intellectual property rights for *howRU*.

Funding

No external funding or support.

Notes

Those wishing to use the EQ-5D should contact the EuroQol Business Management, PO Box 4443, 3006 AK Rotterdam, The Netherlands; <http://www.euroqol.org/>.

Those wishing to use the *howRU* should contact Routine Health Outcomes Ltd. (tim.benson@howrwe.com).

Acknowledgement

The authors are grateful to Dr. Steve Sizmur at the Picker Institute Europe for his suggestions and to the patients and staff at the Whittington Hospital for helping with the data collection.

Ethical approval

Approved by the Joint UCL/UCLH Committee on the Ethics of Human Research (Committee Alpha) Ref 08/H0715/94.

REFERENCES

- Black N. Patient reported outcome measures could help transform healthcare. *Brit Med J* 2013;346:f167. doi: <http://dx.doi.org/10.1136/bmj.f167>.
- The NHS Outcomes Framework 2013–14. London: Department of Health. November 2012.
- Cambria E, Benson T, Eckl C and Hussain A. Sentic PROMs: Application of sentic computing to the development of a novel unified framework for measuring health-care quality. *Expert Syst Appl* 2012;39:10533–43. doi: <http://dx.doi.org/10.1016/j.eswa.2012.02.120>.
- Benson T, Whatling J, Arikian S, Sizmur S, McDonald D and Ingram D. Evaluation of a new short generic measure of health status: *howRU*. *Inform Prim Care* 2010;18:89–101.
- Brooks R, Rabin R and de Charro F. (eds.). *The measurement and valuation of health status using EQ-5D: A European perspective*. Kluwer: Dordrecht, 2003. doi: <http://dx.doi.org/10.1007/978-94-017-0233-1>.
- Szende A, Oppe M and Devlin N (eds.) *EQ-5D value sets: inventory, comparative review and user guide*. Euroqol Group Monographs Volume 2, Dordrecht: Springer, 2007.
- Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095–108. doi: <http://dx.doi.org/10.1097/00005650-199711000-00002>.
- DuBay W. *The principles of readability*. Costa Mesa CA: Impact Information, 2004.
- McDowell I. *Measuring health: A guide to rating scales and questionnaires*. 3rd Edition. New York: Oxford University Press, 2006.
- Paz S, Jiu H, Fongwa M, Morales L and Hays R. Readability estimates for commonly used health-related quality of life surveys. *Qual Life Res* 2009;18:889–900. doi: <http://dx.doi.org/10.1007/s11136-009-9506-y>.
- Hawthorne G, Richardson J, Atherton Day N. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. *Ann of Med* 2001;33:358–70. doi: <http://dx.doi.org/10.3109/07853890109002090>.
- Feng Y, Parkin D and Devlin N. Assessing the performance of the EQ-VAS in the NHS PROMs programme. *Qual Life Res* 2013. doi: <http://dx.doi.org/10.1007/s11136-013-0537-z>.
- Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonse G and Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36. doi: <http://dx.doi.org/10.1007/s11136-011-9903-x>.
- Devlin N and Krabbe P. The development of new research methods for the valuation of EQ-5D-5L. *Eur J Health Econ* 2013;24(Suppl):S1–3.