

Evaluation of Health-Related Quality of Life Measures

Sang-Cheol Bae, MD, PhD, MPH
Department of Internal Medicine
The Hospital for Rheumatic Diseases
Hanyang University, Seoul, Korea

Scope

- **Development & Evaluation of HRQoL**
 - **Reliability**
 - **Validity**
 - **Sensitivity**
 - **Responsiveness**
 - **MCID**
- **Cross-cultural adaptation and validation**

HRQoL development

- **Setting measurement goals**
- **Generating items**
 - literature, targeted patients and experts
- **Forming instrument scales**
 - psychometric analyses
- **Determining scoring methods**
 - sum-score, weighted score
- **Pilot-testing**
 - focus group discussion

HRQoL Evaluation

1. Reliability

degree to which an instrument yields reproducible or consistent scores each time it is administered (same people get same scores)

2. Validity

degree to which an instrument measures what it is supposed (different people get different scores)

HRQoL Evaluation

3. Sensitivity

- ability of an instrument to differentiate between respondents who differ in the attribute to be measured

4. Responsiveness

- ability of an instrument to detect changes in individual respondents over time
- sensitivity to change in longitudinal studies

HRQoL Evaluation

5. Minimal Clinically Important Difference (MCID)

- The smallest change of scores that are subjectively meaningful to patients
- Change large enough for a clinician to base treatment decisions upon it

Two essential elements

: Indicator of minimal change (anchor)

: Amount of HRQOL change among those determined to have changed on anchor

Assessing Reliability

- **Test-retest reliability**

- degree to which an instrument yields repeatable scores if it is administered at different points of time to subjects with stable status
- ICC, Cohen's Kappa

- **Internal consistency reliability**

- degree to which items of a scale are interrelated
- Cronbach's alpha

- **Inter-rater reliability**

- both rater selection and intra-individual response variability
- *kappa* statistic

Assessing Validity

- **Content and face validity**
 - degree to which an instrument's content is representative and relevant
 - degree to which items of an instrument appear to address intended topics clearly and unambiguously
- **Construct validity**
 - degree to which an instrument measures the construct it is designed to measure
 - *a-priori* hypothesis
 - convergent, discriminant, known-groups, criterion validity

Assessing Sensitivity & Responsiveness

Responsiveness (longitudinal validity)

- **Cohen's Effect Size (ES):** mean change / SD(x1)
- **Standardized Response Mean (SRM):** mean change / SD (x2-x1) (SD of intra-subject score changes)
- **Guyatt's Responsiveness (GRI):** mean change / SD (x2-x1) (SD of intra-subject score changes for subjects with stable conditions)
- **Relative Efficiency (RE):** ratio of the squares of the t-statistics for 2 scales

Assessing MCID

- **Example of anchor: People who report a “minimal” change**
 - **How is your physical health now compared to 4 weeks ago?**
 - Much improved**
 - Moderately improved**
 - Minimally improved**
 - No change**
 - Minimally worse**
 - Moderately worse**
 - Much worse**
- **Estimate change in HRQOL among those with minimal change on anchor**

MCID for the CHAQ (0 to 3)

- 136 parents of JRA patients, asked to indicate change in CHAQ that would be “just enough to make a difference”
- Using this anchor based method:
 - MCID improvement: 0.13 reduction in scores
 - MCID deterioration: 0.75 increase in scores

Health-Related Quality of Life

- **Most of measures developed in English & intended for use in English speaking countries**
- **Growing number of multicenter, multinational clinical trials**
- **Need for measures to be used in non English speaking countries & also among immigrant populations**
- **Cultural groups vary in disease expression & use various health care systems**

**“Culture does more than shape illness
as an experience,**

**It shapes the very way we conceive
the illness and health related states”**

Bury, M. Sociol. Health Illness, 1991;13:451-68

How to use HRQoL in non-English speaking countries

- **To develop a new measure in different countries**
: time consuming, not comparable internationally
- **To use a previously developed measure in another language**
: simple translation ?
: need systematic approach for translation & cross-cultural adaptation

Translation

- **Adapt instruments developed and tested in one country and accommodate changes (may require extensive change)**
 - SF36, Nottingham Health Profile, Sickness Impact Profile
 - Forward and backward translation
 - Conceptual, semantic equivalence
- **Simultaneous development of Country Specific Questionnaires and Identify Common Elements (e.g., WHOQOL)**

Cross-cultural development of instrument in another culture

- **Cross-cultural adaptation**
 - : translation**
 - : adjustment of cultural words, idioms, context**
 - : complete transformation of some items**
- **Validation of transformed instrument**
 - : reliability, validity, responsiveness**

Translation Steps

- 1. Author Permission**
- 2. Forward Translation**
- 3. Backward Translation**
- 4. Review of back translation with committee**
- 5. Pilot Study: focus Groups**
- 6. Field Testing**

Translation Steps

1. Author permission

: obtain an approval for translation from the author

2. Forward Translation

: translate instructions, items, and response choices independently by at least two trained bilingual translators

: translators should compare their translations and reconcile discrepancies

Translation Steps

3. Backward Translation

: final forward translation should be back-translated into the original language of the HRQoL instrument by at least two other bilingual translators

: these two translators should compare their backward translations and come to an agreement about any discrepancies

4. Review of back translation with committee

Translation Steps

5. Pilot Study

- : resulting translation should then be cognitively tested in a small sample (about 15) of patients**


- : focus groups**

6. Field Testing

- : the translated instrument administered to a sample of 75 people with related disease**

- : scale equivalence assessed by performing standard reliability and validity testing and comparing these results to those obtained for sample of the original instrument**

- : final report**



The Hospital for Rheumatic Diseases
Section for Clinical & Genetic Epidemiology and Economics
Hanyang University, Seoul, Korea

Thank you!!



Cross-cultural adaptation and validation of the Korean version of the EQ-5D in patients with rheumatic diseases

**The Hospital for Rheumatic Diseases,
Hanyang University, Seoul, South
Korea**

Sang-Cheol Bae, MD, PhD, MPH

Introduction

- **EQ-5D**
 - **Utility-based, generic instrument**
 - **Only 5 items and a VAS**
 - **One of the most widely used instruments for HRQOL**
 - **Korean version – not available**
- **Objectives**
 - **Translation & cross-cultural adaptation of the EQ-5D into Korean**
 - **Evaluation of reliability & validity among various rheumatic patients**

Cover page (1)

EQ - 5D

Health Questionnaire

English version for the UK
(validated for use in Eire)

The descriptive system (2)

By placing a tick in one box in each group below, please indicate which statements best describe your own health state today.

Mobility

- I have no problems in walking about
- I have some problems in walking about
- I am confined to bed

Self-Care

- I have no problems with self-care
- I have some problems washing or dressing myself
- I am unable to wash or dress myself

Usual Activities (e.g. work, study, housework, family or leisure activities)

- I have no problems with performing my usual activities
- I have some problems with performing my usual activities
- I am unable to perform my usual activities

Pain/Discomfort

- I have no pain or discomfort
- I have moderate pain or discomfort
- I have extreme pain or discomfort

Anxiety/Depression

- I am not anxious or depressed
- I am moderately anxious or depressed
- I am extremely anxious or depressed

The EQ VAS (3)

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

Your own
health state
today

The EQ SDQ (4)

Because all replies are anonymous, it will help us to understand your answers better if we have a little background data from everyone, as covered in the following questions.

1. Have you experienced serious illness?
in you yourself
in your family
in caring for others

Yes	No
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

PLEASE TICK APPROPRIATE BOXES
2. What is your age in years?
3. Are you:

Male	Female
<input type="checkbox"/>	<input type="checkbox"/>

PLEASE TICK APPROPRIATE BOX
4. Are you:

a current smoker	<input type="checkbox"/>
an ex-smoker	<input type="checkbox"/>
a never smoker	<input type="checkbox"/>

PLEASE TICK APPROPRIATE BOX
5. Do you now, or did you ever, work in health or social services?

Yes	No
<input type="checkbox"/>	<input type="checkbox"/>

PLEASE TICK APPROPRIATE BOX

If so, in what capacity?
6. Which of the following best describes your main activity?

<i>in employment or self employment</i>	<input type="checkbox"/>
<i>retired</i>	<input type="checkbox"/>
<i>housework</i>	<input type="checkbox"/>
<i>student</i>	<input type="checkbox"/>
<i>seeking work</i>	<input type="checkbox"/>
<i>other (please specify)</i>	<input type="checkbox"/>

PLEASE TICK APPROPRIATE BOX
7. Did your education continue after the minimum school leaving age?

Yes	No
<input type="checkbox"/>	<input type="checkbox"/>

PLEASE TICK APPROPRIATE BOX
8. Do you have a Degree or equivalent professional qualification?

Yes	No
<input type="checkbox"/>	<input type="checkbox"/>

PLEASE TICK APPROPRIATE BOX
9. If you know your postcode, would you please write it here

Methods & Materials (1)

- **Translation and cross-cultural adaptation**
 - According to the EuroQol-Group's guidelines
- **Translation**
 - Forward translation by 2 independent translators
 - Backward translation by another 2 translators
- **Cognitive debriefing**
 - 8 laypersons; 3 rheumatic patients and 5 healthy persons

Methods & Materials (2)

- **Validation**
 - **Construct validity and sensitivity by testing a-priori hypotheses**
 - **Higher EQ-5D scores would have**
 - (Hypothesis 1) Lower scores of condition-specific measures**
 - (Hypothesis 2) Lower score of the 1st question of SF-36 (SF-1)**
 - (Hypothesis 3) Better functional class**
 - (Hypothesis 4) More problems in the physical dimensions of**
the EQ-5D, higher scores of physical scales among condition-specific measures

Methods & Materials (3)

- **Validation study**
 - **Study subjects**
 - **508 patients**
 - **100** **RA**
 - **103** **OA**
 - **111** **SLE**
 - **104** **FMS**
 - **90** **AS**

Methods & Materials (4)

- **Reliability study**
 - **Study subjects**
 - **45 out of consecutive 57 patients with OA or RA at Hanyang Univ Hospital: repeated measure at 1 week interval**
 - **Test-retest reliability**
 - **Intra-class correlation coefficient (ICC)**
 - **Kappa statistics**

Methods & Materials (5)

- **Responsiveness**

- **Study subjects**

- **65 patients with RA : repeated measure at 12 week interval within the context of clinical trial of TNF- α blocker, etanercept[®](before & after treatment)**

- **Internal responsiveness**

- **Paired t-statistics**
 - **Effect size**
 - **Standardized response mean (SRM)**



RESULTS

Translation and adaptation

—Only minor modification

—“I have no problems with self-care”

→ “I have no problems with washing or dressing myself”

—The instruction box to be linked with the thermometer scale

→ black-colored one

General characteristics of the study subjects for validation (I)

Subject Characteristic (unit)	Disease group					Total
	AS	FMS	OA	RA	SLE	
Number of subjects	90	104	103	100	111	508
Gender, <i>female</i> <i>n</i> (%)	12 (13.3)	99 (95.2)	99 (96.1)	93 (93.0)	106 (95.5)	409 (80.5)
Age, <i>years</i> mean (\pm SD)	28.4 (\pm 7.9)	49.3 (\pm 11.1)	61.2 (\pm 6.7)	51.7 (\pm 9.9)	33.2 (\pm 9.9)	45.0 (\pm 15.2)
Education, <i>years</i> mean (\pm SD)	13.8 (\pm 2.2)	10.1 (\pm 3.9)	6.1 (\pm 3.5)	9.1 (\pm 3.9)	13.1 (\pm 2.5)	10.4 (\pm 4.3)
Co-morbidity, <i>yes</i> <i>n</i> (%)	27 (30.0)	21 (20.2)	42 (40.8)	28 (28.0)	60 (54.1)	178 (35.0)

* $p < 0.01$, ** $p < 0.001$, n.s.: not significant by chi-square test or Kruskal-Wallis test in group comparison

General characteristics of the study subjects for validation (II)

	Disease group					Total
	AS	FMS	OA	RA	SLE	
Descriptive statistics of EQ-5D and SF-1						
Dimension of EQ-5D (% of patients with any problems)						
Mobility	47.8	54.8	72.8	64.0	33.3	54.3 **
Self-care	27.8	30.8	28.2	33.0	16.2	27.0 n.s.
Usual activity	61.1	76.0	77.7	70.0	37.8	64.2 **
Pain/discomfort	87.8	93.3	92.2	80.0	65.8	83.5 **
Anxiety/depression	50.0	72.1	60.2	54.0	49.6	56.3 *
EQ-5D _{index} [median, interquartile range (IQR)]						
	0.69 (0.59–0.80)	0.62 (0.52–0.73)	0.62 (0.52–0.69)	0.63 (0.52–0.76)	0.73 (0.66–0.85)	0.69 ** (0.52–0.76)
EQ-5D _{VAS} (median, IQR)						
	60 (40–70)	60 (40–70)	60 (50–70)	70 (50–80)	70 (50–80)	60 ** (50–79)
SF-1 (median, IQR)						
	4 (3–4)	4 (4–5)	4 (3–4)	4 (3–4)	3 (3–4)	4 ** (3–4)

* $p < 0.01$, ** $p < 0.001$, n.s.: not significant by chi-square test or Kruskal-Wallis test in group comparison

Spearman correlation coefficients of the EQ-5D with condition-specific measures and the SF-1 (hypotheses 1 and 2)

Disease	Measure	EQ-5D	
		EQ-5D _{index}	EQ-5D _{VAS}
AS	BASFI	-0.634**	-0.511**
FMS	FIQ	-0.662**	-0.550**
OA	WOMAC – pain	-0.419**	-0.469**
	WOMAC – stiffness	-0.324**	-0.410**
	WOMAC – function	-0.477**	-0.462**
RA	HAQ	-0.608**	-0.389**
SLE	SLEDAI-2K	-0.026 n.s.	0.014 n.s.
	SDI	-0.068 n.s.	-0.288*
Total	SF-1	-0.510**	-0.518**

* $p < 0.01$, ** $p < 0.001$, n.s.: not significant

The EQ-5D scores across the functional classes (hypothesis 3)

Disease	Functional class	EQ-5D	
		EQ-5D _{index}	EQ-5D _{VAS}
OA	I (n=15)	0.73 (0.62–1.00)	70.0 (50.0–80.0)
	II (n=22)	0.62 (0.62–0.69)	60.0 (50.0–70.0)
	III (n=43)	0.62 (0.59–0.69)	60.0 (50.0–70.0)
	IV (n=23)	0.52 (0.12–0.66)	50.0 (40.0–70.0)
	<i>p</i>	0.0002	0.0894
RA	I (n=14)	0.73 (0.73–0.85)	75.0 (70.0–80.0)
	II (n=25)	0.69 (0.62–0.73)	70.0 (50.0–80.0)
	III (n=32)	0.63 (0.55–0.75)	65.0 (45.0–70.0)
	IV (n=28)	0.52 (–0.01–0.62)	55.0 (50.0–70.0)
	<i>p</i>	0.0001	0.0513

Unit: median (IQR)

Association between physical dimensions of the EQ-5D and physical scales among condition-specific measures (hypothesis 4)

Dimension of EQ-5D		AS [BASFI]	OA [WOMAC – function]	RA [HAQ]
Mobility	No problem	0.85 (0.23–1.96)	9.5 (4.0–16.0)	0.50 (0.06–1.00)
	Any problem	3.77 (1.91–5.38)	23.0 (18.0–29.0)	0.88 (0.50–1.50)
	<i>p</i>	<0.0001	<0.0001	0.0009
Self-care	No problem	1.55 (0.33–2.43)	18.0 (11.0–27.0)	0.50 (1.23–0.88)
	Any problem	4.10 (2.59–5.74)	23.0 (18.0–33.0)	1.50 (1.00–1.75)
	<i>p</i>	<0.0001	0.0067	<0.0001
Usual activity	No problem	0.56 (0.13–2.01)	8.0 (4.0–14.0)	0.25 (0.00–0.75)
	Any problem	2.76 (1.63–4.94)	22.0 (17.0–29.0)	0.94 (0.63–1.50)
	<i>p</i>	<0.0001	<0.0001	<0.0001

Unit: median (IQR)

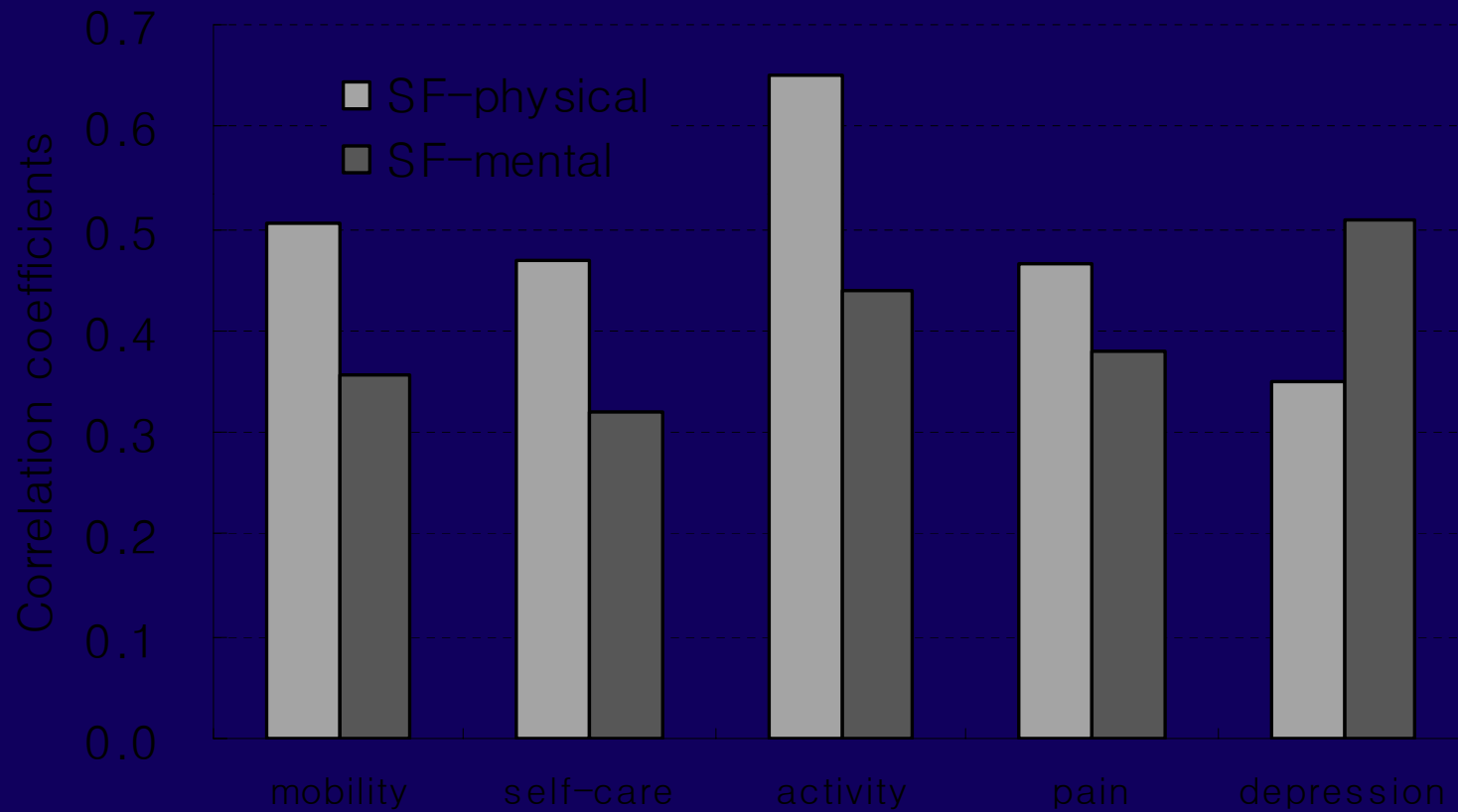
Reliability of the EQ-5D between repeated measures at 1-week interval

1) Test-retest reliability			
EQ-5D scores	Test median (IQR)	Re-test median (IQR)	ICC (95% CI)
EQ-5D _{index}	0.62 (0.34–0.73)	0.62 (0.39–0.69)	0.751 (0.590–0.855)
EQ-5D _{VAS}	60.0 (50.0–75.0)	70.0 (50.0–77.5)	0.767 (0.615–0.865)
2) Agreement			
Dimensions of EQ-5D	Kappa statistics (95% CI)		<i>p</i>
Mobility	0.665 (0.439–0.892)		0.706
Self-care	0.527 (0.271–0.784)		1.000
Usual activity	0.690 (0.461–0.918)		0.414
Pain/discomfort	0.455 (0.010–0.900)		0.317
Anxiety/depression	0.772 (0.585–0.959)		0.180

Responsiveness of the EQ-5D_{index} and EQ-5D_{VAS} between pre- and post-treatment

Attribute	EQ-5D _{index}	EQ-5D _{VAS}
Pretreatment, mean (\pm SD)	0.38 (\pm 0.33)	54.8 (\pm 19.9)
Post-treatment, mean (\pm SD)	0.59 (\pm 0.19)	65.8 (\pm 19.0)
Difference, mean (\pm SD)	0.22 (\pm 0.33)	10.9 (\pm 26.6)
Paired <i>t</i> statistic (<i>p</i>)	4.94 (<0.0001)	3.09 (0.0031)
Effect size	0.658	0.548
Standardized response mean	0.649	0.410

Convergent & discriminant validity



Summary

- **The Korean EQ-5D**
 - **Reliability : good**
 - **Validity & sensitivity : good**
 - **Responsiveness : good**

Conclusion

- **KEQ-5D : applicable to Korean patients with various rheumatic diseases**