Health-Related Quality of Life (HRQOL) - 2014

John E. Ware, Jr., PhD

Professor & Division Chief, Dept. of Quantitative Health Sciences, UMass Medical School
Chief Science Officer and Founder, JWRG, Incorporated, Worcester, MA

Introduction to the Principles and Practice of Clinical Research
National Institutes of Health - Warren G. Magnuson Clinical Center Building 10
Lipsett Amphitheater, Bethesda, MD, December 2, 2013
“The best measure of quality is not how well or how frequently a medical service is given, but how closely the result approaches the fundamental objectives of prolonging life, relieving distress, restoring function, and preventing disability.”

Lembcke, 1952
Quality of Life (QoL)

- Community
- Education
- Family Life
- Friendships
  - Health
- Housing
- Marriage
- Nation
- Neighborhood
- Self
- Standard of Living
- Work

Source: Campbell, 1981
World Health Organization
Definition of Health

“Health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity”

WHO, 1948
Health is Measured in Terms of:

• **Bodily** structure & function

• Specific **symptoms**

• What you do/are able to do – **functioning**

• How you feel – **subjective** ill- and well-being (+ and -)

• What you say it is – **personal** evaluation
Continuum of Disease-specific and Generic Health Measures - Arthritis

X-ray, Disease Progression

Clinical Markers

(1)

Arthritis Symptoms

Have you had the following symptoms:
Joint pain, swelling, burning sensation:
• Almost every day
• Several days a week
• A few days a month
• Not at all

(2)

Arthritis Impact

How much does your arthritis limit your usual activities or enjoyment of everyday life?
• Not at all
• A little
• Moderately
• Extremely

(3)

Arthritis Symptoms

Generic Functioning, Well-being and Evaluation

(4)

In general, would you say your health is:
• Excellent
• Very good
• Good
• Fair
• Poor

Health-related QOL (HRQoL)

Adapted from: Wilson and Cleary, JAMA, 1995
Ware, Annual Rev. Pub. Health, 1995
There is More to the Continuum

1. Clinical Markers
2. Specific Symptoms
3. Impact of Disease-specific Problems
4. Generic Functioning, Well-being and Evaluation
Predictive Validity: HRQoL is One of the Best Predictors

Health-Related QOL (HRQoL)

Impact of Disease-specific Problems

Generic Functioning, Well-being and Evaluation

(3) (4)

Future health
Inpatient expenditures
Outpatient expenditures
Job loss
Response to treatment
Return to work
Work productivity
Mortality
## Summary of Content/Concepts for Widely-Used Generic Health Surveys

**Source:** Adapted from Ware, 1995

### Psychometric Utility Concepts and Characteristics

<table>
<thead>
<tr>
<th>Concepts and Characteristics</th>
<th>SIP</th>
<th>HIE</th>
<th>NHP</th>
<th>QLI</th>
<th>COOP</th>
<th>DUKE</th>
<th>MOS FWBP</th>
<th>MOS SF-36</th>
<th>PROMIS</th>
<th>QWB</th>
<th>EQ-5D</th>
<th>HUI</th>
<th>SF-6D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONCEPTS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical functioning (-)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Social functioning (-)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Role functioning (-)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Psychological distress (-)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Health perception (gen’l)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Pain (-)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Vitality (-)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Psychological well-being</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Sleep</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Cognitive functioning</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Quality of life</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Reported health transition</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

**SIP** = Sickness Impact Profile (1976)
**HIE** = Health Insurance Experiment surveys (1979)
**NHP** = Nottingham Health Profile (1980)
**QLI** = Quality of Life Index (1981)
**COOP** = Dartmouth Function Charts (1987)
**DUKE** = Duke Health Profile (1990)
**MOS FWBP** = MOS Functioning & Well-Being Profile (1992)
**MOS SF-36** = MOS 36-Item Short-Form Health Survey (1992)
**QWB** = Quality of Well-Being Scale (1973)
**EUROQOL** = European Quality of Life Index (1990)
**HUI** = Health Utility Index (1996)
**SF-6D** = SF-36 Utility Index (2002)

Source: Adapted from Ware, 1995
PROMIS Website
www.nihpromis.org

PROMIS References


Medical Outcomes Study
36-Item Health Survey (SF-36)

Items

3a. Vigorous Activities
3b. Moderate Activities
3c. Lift, Carry Groceries
3d. Climb Several Flights
3e. Climb One Flight
3f. Bend, Kneel
3g. Walk Mile
3h. Walk Several Blocks
3i. Walk One Block
3j. Bathe, Dress
4a. Cut Down Time
4b. Accomplished Less
4c. Limited in Kind
4d. Had Difficulty
7. Pain-Magnitude
8. Pain-Interfere
1. EVGFP Rating
11a. Sick Easier
11b. As Healthy
11c. Health To Get Worse
11d. Health Excellent
9a. Pep/Life
8e. Energy
9g. Worn Out
9f. Tired
5. Social-Extent
10. Social-Time
5a. Cut Down Time
5b. Accomplished Less
5c. Not Careful
9b. Nervous
9c. Down in Dumps
9d. Peaceful
9f. Blue/Sad
9h. Happy

Scales

Physical Functioning (PF)
Role-Physical (RP)
Bodily Pain (BP)
General Health (GH)*
Vitality (VT)*
Social Functioning (SF)*
Role-Emotional (RE)
Mental Health (MH)

Summary Measures

Physical

Mental

Utility Index
(Brazier et al., 2002)
(Lam, Brazier, McGhee 2008)

* Significant correlation with other summary measure.
Generic Health Profiles: Before & After Medication

<table>
<thead>
<tr>
<th>Component Summaries</th>
<th>SF-36 Subscales</th>
<th>SF-36 Subscales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthritis</td>
<td></td>
<td>Depression</td>
</tr>
<tr>
<td>Norm</td>
<td>Physical</td>
<td>Physical</td>
</tr>
<tr>
<td>Best Health</td>
<td>Mental</td>
<td>Mental</td>
</tr>
<tr>
<td>Poorest Health</td>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SF-36 Subscales</th>
<th>Component Summaries</th>
<th>Treatment</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical</td>
<td></td>
<td>P</td>
<td>M</td>
</tr>
<tr>
<td>Mental</td>
<td></td>
<td>P</td>
<td>M</td>
</tr>
</tbody>
</table>

12
Next Step: Integrate and Standardize Disease-specific and Generic Measures

Example: Osteoarthritis
# Comparison of Content of Generic And Disease-Specific Measures

<table>
<thead>
<tr>
<th>Health Domains</th>
<th>Generic*</th>
<th>Disease-Specific*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SIP</td>
<td>MOS FWBP</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Social functioning</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Role functioning</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Psychological distress</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Health perceptions (general)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Pain (bodily)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Energy/fatigue</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Psychological well-being</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Sleep</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Cognitive functioning</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Quality of life</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Health transition (B/S/W)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Sexual functioning</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Specific symptoms</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Specific treatments</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

SIP - Sickness Impact Profile, 136 items (1976)
MOS FWBP - MOS Functioning and Well-Being Profile, 149 items (1992)
PROMIS - Patient Reported Outcomes Measurement Information System (2007-on)

*Source: Comparison adapted from Ware 1987; Ware, 1995

DQOL - Diabetes Quality of Life Measure
KDQOL - Kidney Disease Quality of Life Questionnaire
MLHQF - Minnesota Living with Heart Failure Questionnaire
SAQ - Seattle Angina Questionnaire
St. George's - Respiratory Questionnaire
WOMAC - Western Ontario and Mac Master’s Arthritis Index
Content Areas*
QOL Disease Impact Scale (QDIS)

1. Global (“activity” and “QOL”)
2. Physical Functioning
3. Mobility
4. General Health
5. Fatigue
6. Cognitive
7. Emotional
8. Sleep
9. Role Functioning
10. Social Functioning

*Note: Bank of 49 items reduced to 25 items with attribution to specific condition
Anatomy of a Survey Item:
QDIS Changes Attribution from Health to Specific Disease

During the past 4 weeks, how often did your arthritis limit your ability to do your everyday activities?

- Very often
- Often
- Sometimes
- Rarely
- Never


QDIS Reference: QOL Disease Impact Scale (QDIS); Ware JE, Guyer R, Harrington M, Boulanger R. Quality of Life Research, 2012.
Better Measures Are Being Constructed

- Standardized Metrics
- Adaptive survey administrations
- Norm-based scoring
- Internet and mobile data collection
Improving the Physical Function “Ruler”

Mean = 50
SD = 10

1980
Old “Ruler”
> 75%
@ Ceiling

1990
NEW “Ruler”
> 30%
@ Ceiling

2008
BETTER “Ruler”
< 3 % @ Ceiling

Vigorous Activities, Not limited

Norm

Climbing several flights of stairs
Walk one hundred yards
Bathing or dressing, Limited a little

Mean = 50
SD = 10
We Need the Health Equivalent of a Two-Sided Tape Measure

and Public-Private Partnerships Meeting the Needs of Research and Business
Short-Form Surveys and Ceiling Effects

Measuring Too Low - Ceiling Effect
Some Thermometers Focus on a Very Narrow Range

Cooking Thermometer

130–190 °F
54–88 °C
Example: Cross-Calibrating Celsius and Fahrenheit

Temperature

Normal Human Blood

Shirt Sleeve Weather

Water Freezes

F° = 98.6  C° = 37.0
Solution: Adaptive Survey Methods

Patient scores here

CAT = Computerized Adaptive Testing
Adaptive Assessments of Disease Impact
Match Questions to Each Patient’s Level

Higher is worse.
Mean = 50
SD = 10
First Question
Noisy Score Estimate (+/- 15)

Score estimate, 1st response = 62 +/- 15

Higher is worse. Mean = 50 SD = 10
Second Question: Standard Error Reduced by One Third

Score estimate, 2 responses = 64 +/- 10

Higher is worse. Mean = 50 SD = 10
Third Question: Standard Error Cut in Half

Score estimate, 3 responses = 63 +/- 7

Mean = 50  SD = 10

Higher is worse.
Fourth Question: Standard Error
Cut by Two Thirds

Score estimate, 4 responses = 62 +/- 5

Higher is worse.
Mean = 50
SD = 10
Practical Implications of CAT in Health Assessment

MEDICAL CARE

Volume 38, Number 9, Supplement II, pp II-73–II-82
©2000 Lippincott Williams & Wilkins, Inc.

Practical Implications of Item Response Theory and Computerized Adaptive Testing

A Brief Summary of Ongoing Studies of Widely Used Headache Impact Scales

JOHN E. WARE, JR, PhD, Jakob B. Bjorner, MD, PhD, and Mark Kosinski, MA

We have the potential to substantially advance the field of health status assessment by constructing and calibrating questionnaires based on item response theory (IRT) and administering them using computerized adaptive methods. This opportunity could untreated. It was hoped that an accurate and user-friendly report of headache impact would be useful to patients and those who treat them. To benefit as much as possible from prior work and to maintain comparability of scores with current
A Promising Solution in 1999: CAT-Based Health Assessment

Performance of 5-item CAT Scores Confirmed in NIH-Sponsored Studies

- **Mental Health**: $r = 0.98$, $N = 2,753$
- **Headache Disability**: $r = 0.94$, $N = 1,016$
- **Pediatric Disability**: $r = 0.96$, $N = 263$
- **Chronic Kidney Disease**: $r = 0.95$, $N = 1,846$
- **Diabetes Impact**: $r = 0.93$, $N = 100$
- **Post Acute Rehabilitation**: $r = 0.95$, $N = 485$
What are the Advantages of Dynamic Assessments?

- More accurate risk screening
- Reliable enough to monitor individual outcomes
- Brevity of a short form – 90% reduction in respondent burden
- Elimination of “ceiling” & “floor” effects
- Can be administered using various data collection technologies
- Markedly reduced data collection costs
- Monitor data quality in real time
Matching Methods to Applications

- Population Surveys
- Group-Level Studies
- Patient-Level Assessments

- Single-Item
- Multi-Item Scale
- “Item Bank” (CAT Dynamic)

- Noisy Individual Classification
- Very Accurate Individual Classification

Most Functionally Impaired
Adaptive Survey Logic (ASLX®) Flow Chart

**Survey Content And User Interface (UI)**
- Estimate Score & CI from Screen
  - ASLX® Screen (NEG)
  - POS
  - Supplemental Domain Assessment (CAT/Static)*

Repeat Cycle for Each Generic Domain

**Generic Health Assessments**
- QDIS Assessments (QOLIX® CCC)

Follow-up Assessment Modules

Condition Present?
- YES
  - ASLX® Screen (Neg)
  - Pos
  - Supplemental Disease Impact Assessment (CAT/Static)*

Repeat Cycle for Each Chronic Condition

- Impact Estimated From Screening
- Next Condition

**Abbreviations:** QOLIX® DIS; Chronic condition checklist QOLIX® CCC), Computer adaptive test (CAT), Confidence interval (CI), Diagnosis (Dx), Electronic data capture (EDC), User interface (UI); Follow-up assessment modules vary across applications.
HRQoL Validation Strategies

Clinical Causes
- Diagnosis
- Disease severity
- Clinical endpoint
- Treatment

Gold Standard

HR-QOL

Other Measures & Methods

Economic & Social Consequences
- Work productivity
- Costs of care
- Mortality
- Self-evaluated health

Interpreting HRQoL Scores

Chronic Disease Registry

Population Health Surveys

Congestive Heart Failure
Chronic Lung Disease
Chronic Kidney Disease
Diabetes Type II
Asthma
Average Adult
Average Well Adult

MOS Physical Component Summary (PCS)
(Mean = 50, SD = 10)
Interpreting HRQoL Scores

- Content of questionnaire items
- Statistically significant change
- Important reduction in disease burden
- Reduction in subsequent expenditures
- Substantial increase in work productivity

**MOS Physical Component Summary (PCS)**
(Mean = 50, SD = 10)

- **Diabetes Type II**
- **Asthma**
- **Congestive Heart Failure**
- **Chronic Kidney Disease**
- **RA Before Rx**
- **After Rx**

**Population Health Surveys**
**Clinical Trials**
**Chronic Disease Registry**

- Strand et al. Arth & Rheum, 1999;42(9): 1870-78
Are Generic HRQoL Measures Responsive?

X-ray, Disease Progression

Clinical Markers

Specific Symptoms

Impact of Disease-specific Problems

Generic Health

SF-36 agreed with primary endpoint (across Rx):
- 219 of 253 RCTs
- 86.6%

In general, would you say your health is:
- Excellent
- Very good
- Good
- Fair
- Poor

Internet Sampling and Data Collection


Health Outcomes Research
Using Handhelds - Mobile Metrics
Final Comment

For more information: john.ware@umassmed.edu

Business Week 11/26/2001