Patient-Centered Outcomes of Health Care

CTSI Training Module 3C
Comparative Effectiveness Research
January 23, 2014
8:30am - 12:30pm
CHS 17-187

Ron D. Hays, Ph. D.
Introduction to Patient-Reported Outcomes

8:30-9:30am
U.S. Health Care Issues

• **Access to care**
  – ~ 50 million people without health insurance

• **Costs of care**
  – Expenditures ~ $ 2.7 Trillion

• **Effectiveness (quality) of care**
How Do We Know If Care Is Effective?

• Effective care maximizes probability of desired health outcomes
  – Health outcome measures indicate whether care is effective

  Cost ↓

  — — — — — — — — — — — — — — — —
  Effectiveness ↑
Health Outcomes Measures

• Traditional clinical endpoints
  – Survival
  – Clinical/biological indicators
    • Rheumatoid factor
    • Blood pressure
    • Hematocrit

• Patient-Reported Outcomes
Patient-Reported Measures (PRMs)

- **Mediators**
  - Health behaviors (adherence)

- **Health Care Process**
  - Reports about care (e.g., communication)

- **Outcomes (PROs)**
  - Patient satisfaction with care
  - Health-Related Quality of Life (HRQOL)
Health-Related Quality of Life (HRQOL)

How the person FEELs (well-being)
- Emotional well-being
- Pain
- Energy

What the person can DO (functioning)
- Self-care
- Role
- Social
HRQOL is Not

Quality of environment
Type of housing
Level of income
Social Support
Types of HRQOL Measures

- Targeted vs. Generic
- Profile vs. Preference-based
Targeted Item

During the last 4 weeks, how often were you angry about your irritable bowel syndrome?

None of the time
A little of the time
Some of the time
Most of the time
All of the time
Burden of Kidney Disease (Targeted Scale)

- My kidney disease interferes too much with my life.
- Too much of my time is spent dealing with my kidney disease.
- I feel frustrated with my kidney disease.
- I feel like a burden on my family.
Generic Item

In general, how would you rate your health?

Excellent
Very Good
Good
Fair
Poor
Does your health now limit you in walking more than a mile?

(If so, how much?)

Yes, limited a lot
Yes, limited a little
No, not limited at all
How much of the time during the past 4 weeks have you been happy?

None of the time
A little of the time
Some of the time
Most of the time
All of the time
Generic Profile (SF-36)

- Physical functioning (10 items)
- Role limitations/physical (4 items)
- Role limitations/emotional (3 items)
- Social functioning (2 items)
- Emotional well-being (5 items)
- Energy/fatigue (4 items)
- Pain (2 items)
- General health perceptions (5 items)
Scoring HRQOL Scales

• Average or sum all items in the same scale.

• Transform average or sum to
  • 0 (worse) to 100 (best) possible range
  • z-score (mean = 0, SD = 1)
  • T-score (mean = 50, SD = 10)
Linear Transformations

\[ X = \frac{(\text{original score} - \text{minimum}) \times 100}{(\text{maximum} - \text{minimum})} \]

\[ Y = \text{target mean} + (\text{target SD} \times Z_X) \]

\[ Z_X = \frac{(X - \bar{X})}{SD_x} \]
HRQOL in HIV Compared to other Chronic Illnesses and General Population

- MS
- ESRD
- Diabetes
- Depression
- Prostate disease
- GERD
- Epilepsy
- General Pop
- AIDS
- Symptomatic
- Asymptomatic

T-score metric

Hays et al. (2000), American Journal of Medicine
Physical Health

- Physical function
- Role function physical
- Pain
- General Health
Mental Health

- Emotional Well-Being
- Role function - emotional
- Energy
- Social function
SF-36 PCS and MCS

PCS\_z = \((PF\_Z \times 0.42) + (RP\_Z \times 0.35) + (BP\_Z \times 0.32) + (GH\_Z \times 0.25) + (EF\_Z \times 0.03) + (SF\_Z \times -0.01) + (RE\_Z \times -0.19) + (EW\_Z \times -0.22)\)

MCS\_z = \((PF\_Z \times -0.23) + (RP\_Z \times -0.12) + (BP\_Z \times -0.10) + (GH\_Z \times -0.02) + (EF\_Z \times 0.24) + (SF\_Z \times 0.27) + (RE\_Z \times 0.43) + (EW\_Z \times 0.49)\)

PCS = \((PCS\_z \times 10) + 50\)

MCS = \((MCS\_z \times 10) + 50\)
536 Primary Care Patients Initiating Antidepressant Tx

- 3-month improvements in physical functioning, role-physical, pain, and general health perceptions ranging from 0.28 to 0.49 SDs.
  - Trivial < 0.20 SD
  - Small = 0.20 SD
  - Medium = 0.50 SD
  - Large = 0.80 SD
- Yet SF-36 PCS did not improve.

*Simon et al. (Med Care, 1998)*
Lower scores than general population on
- Emotional well-being (↓ 0.3 SD)
- Role—emotional (↓ 0.7 SD)
- Energy (↓ 1.0 SD)
- Social functioning (↓ 1.0 SD)

Yet SF-36 MCS was only 0.2 SD lower.

Nortvedt et al. (Med Care, 2000)
Farivar et al. alternative weights

\[ \text{PCS}_z = (\text{PF}_z \times 0.20) + (\text{RP}_z \times 0.31) + (\text{BP}_z \times 0.23) + (\text{GH}_z \times 0.20) + (\text{EF}_z \times 0.13) + (\text{SF}_z \times 0.11) + (\text{RE}_z \times 0.03) + (\text{EW}_z \times -0.03) \]

\[ \text{MCS}_z = (\text{PF}_z \times -0.02) + (\text{RP}_z \times 0.03) + (\text{BP}_z \times 0.04) + (\text{GH}_z \times 0.10) + (\text{EF}_z \times 0.29) + (\text{SF}_z \times 0.14) + (\text{RE}_z \times 0.20) + (\text{EW}_z \times 0.35) \]

Is New Treatment (X) Better Than Standard Care (O)?

Physical Health
X > 0

Mental Health
0 > X
## Is Medicine Related to Worse HRQOL?

<table>
<thead>
<tr>
<th>Person</th>
<th>Medication Use</th>
<th>HRQOL (0-100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>dead</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>dead</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>75</td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>Yes</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Yes</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>Yes</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>Yes</td>
<td>75</td>
</tr>
<tr>
<td>10</td>
<td>Yes</td>
<td>100</td>
</tr>
</tbody>
</table>

### Group

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>HRQOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Medicine</td>
<td>3</td>
<td>75</td>
</tr>
<tr>
<td>Yes Medicine</td>
<td>5</td>
<td>50</td>
</tr>
</tbody>
</table>
Quality of Life for Individual Over Time

quality of life (HRQL)

years

now

death

1.0

0
Ultimate Use of HRQOL Measures--Helping to Ensure Access to Cost-Effective Care

Cost ↓

Effectiveness (“Utility”) ↑
http://www.ukmi.nhs.uk/Research/pharma_res.asp
What is value?
- Preference or desirability of health states

How are QALYs used?
- Societal resource allocation
- Personal decisions such as decision about whether to have a treatment
- Societal or program audit
  - Evaluate programs in terms of health of the population.
Direct Preference Measures

- Underlying attributes unknown
  - Rating Scale
  - Standard gamble
  - Time tradeoff
Overall, how would you rate your current health?
(Circle One Number)

0       1       2        3       4       5        6      7       8        9     10

Worst possible health (as bad or worse than being dead)
Half-way between worst and best
Best possible health
Standard Gamble

SURE THING

GAMBLE

Lifetime of Back Pain (U = ?)

No Pain or Disability (U = 1.0)

1 - p

Perioperative Death (U = 0)
Time Tradeoff

Alternative 1 is current health for time “t” (given), followed by death. Alternative 2 is full health for time “x” (elicited), followed by death. 

\[ \frac{x}{t} = \text{preference for current health} \]
Utility Assessments

An important issue in medical decision making is how to measure people's preferences for health states in a way that will facilitate comparisons of health states. The most important measure of preference is the "utility" of the health state to the individual who will experience it, which is a value from 0 (representing death) to 1 (perfect health and well-being).

This page allows you to assess the utility for a health state using three techniques: rating scale, standard gamble, and time tradeoff.

Enter the health state that you'd like to assess the utility of: amputation of your left hand at the wrist

Select the assessment method to use:

- Rating scale
- Standard Gamble
- Time Tradeoff

Let's do it!

http://araw.mede.uic.edu/cgi-bin/utility.cgi
SG > TTO > RS

- SG = TTO\(^a\)
- SG = RS\(^b\)

Where \(a\) and \(b\) are less than 1
Indirect Preference Measures

- Attributes know
- Based on “societal preferences” a single score is assigned
  - Quality of Well-Being (QWB) Scale
  - EQ-5D
  - HUI2 and HUI3
  - SF-6D
Quality of Well-Being (QWB) Scale

- Summarize HRQOL in QALYs
  - Mobility (MOB)
  - Physical activity (PAC)
  - Social activity (SAC)
  - Symptom/problem complexes (SPC)

Well-Being Formula: \( w = 1 + MOB + PAC + SAC + SPC \)
Quality of Well-Being Weighting Procedure

Each page in this booklet tells how an imaginary person is affected by a health problem on one day of his or her life. I want you to look at each health situation and rate it on a ladder with steps numbered from zero to ten.

The information on each page tells 1) the person's age group, 2) whether the person could drive or use public transportation, 3) how well the person could walk, 4) how well the person could perform the activities usual for his or her age, and 5) what symptom or problem was bothering the person.

Adult (18-65)
Drove car or used public transportation without help (MOB)
Walked without physical problems (PAC)
Limited in amount or kind of work, school, or housework (SAC)
Problem with being overweight or underweight (SYM)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Perfect Health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>Death</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
EQ-5D (243 states, 3 levels each)

- Mobility
- Self-care
- Usual activities
- Pain/discomfort
- Anxiety/depression

http://www.euroqol.org/
Your own health state today

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

Do not tick more than one box in each group.

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 and 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
SF-6D

Brazier et al. (1998, 2002)

- 6-dimensional classification (collapsed role scales, dropped general health)
- Uses 11 SF-36 items (8 SF-12 and 3 additional physical functioning items)

--- 18,000 possible states

--- 249 states rated by sample of 836 from UK general population

http://www.shef.ac.uk/scharr/sections/heds/mvh/sf-6d
Health state 424421 (0.59)

• Your health limits you a lot in moderate activities (such as moving a table, pushing a vacuum cleaner, bowling or playing golf)
• You are limited in the kind of work or other activities as a result of your physical health
• Your health limits your social activities (like visiting friends, relatives etc.) most of the time.
• You have pain that interferes with your normal work (both outside the home and housework) moderately
• You feel tense or downhearted and low a little of the time.
• You have a lot of energy all of the time
# Correlations Among Indirect Measures

<table>
<thead>
<tr>
<th></th>
<th>EQ-5D</th>
<th>HUI2</th>
<th>HUI3</th>
<th>QWB-SA</th>
<th>SF-6D</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ-5D</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HUI2</td>
<td>0.71</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HUI3</td>
<td>0.68</td>
<td>0.89</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QWB</td>
<td>0.64</td>
<td>0.66</td>
<td>0.66</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>SF-6D</td>
<td>0.70</td>
<td>0.71</td>
<td>0.69</td>
<td>0.65</td>
<td>1.00</td>
</tr>
</tbody>
</table>

## Change in Indirect Preference Measures Over Time

<table>
<thead>
<tr>
<th></th>
<th>Cataract (1 mon. – B)</th>
<th>Heart F (6 mons. – B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HUI3</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>HUI2</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>QWB-SA</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>EQ-5D</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>SF-6D</td>
<td>0.00</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Kaplan, R. M. et al. (2011). Five preference-based indexes in cataract and heart failure patients were not equally responsive to change. J Clinical Epidemiology, 64, 497-506.

ICC for change was 0.16 for cataract and 0.07 for heart failure.

Break #1
Development and Evaluation of Patient-reported Outcomes

9:45-10:45am

Measurement and Modeling of Health-Related Quality of Life

R.D. Hays, University of California at Los Angeles, Los Angeles, CA, USA
B.B. Reeve, National Cancer Institute, Bethesda, MD, USA
© 2008 Elsevier Inc. All rights reserved.

Introduction

Health-related quality of life (HRQOL) refers to how well a person functions in their life and how or how perceived well-being in physical, mental, and social domains of health. HRQOL includes whether the person can carry out a range of activities of daily living such as bathing or dressing him- or herself (physical functioning). It also includes whether the person can climb stairs, walk, or run. Other relevant aspects of functioning include the extent to which one is able to interact with family, friends, and others (social functioning). The functional part of HRQOL consists of behaviors that can be observed by other people.

The well-being part of HRQOL refers to internal, subjective perceptions such as vitality, pain, anxiety, depressive symptoms, and general health perceptions. These perceptions are not directly observable by others. A person who is anxious might look nervous to an external observer or someone in pain might grimace, but these external signs can be hidden, difficult to detect, and provide at best an indirect indicator of the way the person feels.

Methods of Assessing Health-Related Quality of Life

The target person is considered the best source of information about his or her functioning and well-being. Hence, the usual mode of assessing HRQOL is through self-report. HRQOL data are typically gathered using...
End goal is measure that is “Psychometrically Sound”

- Same people get same scores

- Different people get different scores and differ in the way you expect

- Measure works the same way for different groups (age, gender, race/ethnicity)

- Measure is practical
Measurement Steps

• Review literature
• Focus groups
  – Define constructs and draft items
• Pretest (cognitive interviews)
  – Revise items
• Field test
  – Analyze and finalize items
Focus Groups

• Discuss feelings, attitudes, perceptions
• Learn
  – Vocabulary and thinking patterns
• Conversational meeting
  – Moderator and 6-12 people
  – Questions posed
  – Group synergy
  – Economical
Pretesting

“Cut and try, see how it looks and sounds, see how people react to it, and then cut again, and try again” Converse & Presser (1986, p. 78)

Identify problems with

- Comprehension of items (stem/response options)
- Retrieval of information
- Skip patterns
- Response burden
Cognitive Interviews

- “Think aloud”
- Intermittent probes
- Retrospective recall
Flesch-Kincaid Grade Level

FK GL = \(0.39\) * (n of words/n of sentences) + \(11.8\) * (n of syllables/n of words) – 15.59

• Driven by sentence length and syllables per word

• U.S. school grade level (e.g., 8.0 implies that 8th grader can understand the document).

• Possible minimum = -3.4
  – Green eggs and ham averages 5.7 words per sentence and 1 syllable per word
  – (FK GL = -1.3)
Listed below are a few statements about your relationships with others. How much is each statement TRUE or FALSE for you

<table>
<thead>
<tr>
<th>Statement</th>
<th>Definitely True</th>
<th>Mostly True</th>
<th>Don’t Know</th>
<th>Mostly False</th>
<th>Definitely False</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I am always courteous even to people who are disagreeable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. There have been occasions when I took advantage of someone.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. I sometimes try to get even rather than forgive and forget.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. I sometimes feel resentful when I don’t get my way.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. No matter who I’m talking to, I’m always a good listener.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Scoring Multi-Item Scales

- Average or sum all items in the same scale.

- Transform average or sum to
  - 0 (worse) to 100 (best) possible range
  - z-score (mean = 0, SD = 1)
  - T-score (mean = 50, SD = 10)
Listed below are a few statements about your relationships with others. How much is each statement TRUE or FALSE for you

<table>
<thead>
<tr>
<th>Statement</th>
<th>Definitely True</th>
<th>Mostly True</th>
<th>Don’t Know</th>
<th>Mostly False</th>
<th>Definitely False</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I am always courteous even to people who are disagreeable.</td>
<td>100</td>
<td>75</td>
<td>50</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>2. There have been occasions when I took advantage of someone.</td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>3. I sometimes try to get even rather than forgive and forget.</td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>4. I sometimes feel resentful when I don’t get my way.</td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>5. No matter who I’m talking to, I’m always a good listener.</td>
<td>100</td>
<td>75</td>
<td>50</td>
<td>25</td>
<td>0</td>
</tr>
</tbody>
</table>
Create T-score

z-score = (score – 36)/31
T-score = (10 * z-score) + 50

z-score = (100- 36)/31 = 2.06
T-score = 71
Reliability

• Extent to which measure yields similar result when the thing being measured hasn’t changed

• Ranges from 0-1
## Reliability and Intraclass Correlation

<table>
<thead>
<tr>
<th>Model</th>
<th>Reliability</th>
<th>Intraclass Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>One-way</strong></td>
<td>$\frac{MS_{BMS} - MS_{WMS}}{MS_{BMS}}$</td>
<td>$\frac{MS_{BMS} - MS_{WMS}}{MS_{BMS} + (k - 1)MS_{WMS}}$</td>
</tr>
<tr>
<td><strong>Two-way fixed</strong></td>
<td>$\frac{MS_{BMS} - MS_{EMS}}{MS_{BMS}}$</td>
<td>$\frac{MS_{BMS} - MS_{EMS}}{MS_{BMS} + (k - 1)MS_{EMS}}$</td>
</tr>
<tr>
<td><strong>Two-way random</strong></td>
<td>$\frac{N(MS_{BMS} - MS_{EMS})}{NMS_{BMS} + MS_{JMS} - MS_{EMS}}$</td>
<td>$\frac{MS_{BMS} - MS_{EMS}}{MS_{BMS} + (k - 1)MS_{EMS} + k(MS_{JMS} - MS_{EMS}) / N}$</td>
</tr>
</tbody>
</table>

**BMS** = Between Ratee Mean Square  \hspace{1cm} **N** = n of ratees  
**WMS** = Within Mean Square  \hspace{1cm} **k** = n of items or raters  
**JMS** = Item or Rater Mean Square  
**EMS** = Ratee x Item (Rater) Mean Square
## Two-Way Fixed Effects (Cronbach’s Alpha)

### Source

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondents (BMS)</td>
<td>5</td>
<td>15.67</td>
<td>3.13</td>
</tr>
<tr>
<td>Items (JMS)</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Resp. x Items (EMS)</td>
<td>5</td>
<td>2.00</td>
<td>0.40</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>11</td>
<td>17.67</td>
<td></td>
</tr>
</tbody>
</table>

\[
\text{Alpha} = \frac{3.13 - 0.40}{3.13} = \frac{2.93}{3.13} = 0.87
\]

\[
\text{ICC} = 0.77
\]
Reliability Minimum Standards

• 0.70 or above (for group comparisons)

• 0.90 or higher (for individual assessment)

- SEM = SD (1- reliability)^{1/2}
- 95% CI = true score +/- 1.96 x SEM
  - if true z-score = 0, then CI: -.62 to +.62
  - Width of CI is 1.24 z-score units
Range of reliability estimates

0.80-0.90 for blood pressure
0.70-0.90 for multi-item self-report scales

Category Response Curves

Appreciating each day.

Probability of Response

Posttraumatic Growth

No change
Great change
Very great change

No change
Small change
Moderate change
Great change

Very small change

Appreciating each day.
### Item-scale correlation matrix

<table>
<thead>
<tr>
<th></th>
<th>Depress</th>
<th>Anxiety</th>
<th>Anger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item #1</td>
<td>0.80*</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Item #2</td>
<td>0.80*</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Item #3</td>
<td>0.80*</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Item #4</td>
<td>0.20</td>
<td>0.80*</td>
<td>0.20</td>
</tr>
<tr>
<td>Item #5</td>
<td>0.20</td>
<td>0.80*</td>
<td>0.20</td>
</tr>
<tr>
<td>Item #6</td>
<td>0.20</td>
<td>0.80*</td>
<td>0.20</td>
</tr>
<tr>
<td>Item #7</td>
<td>0.20</td>
<td>0.20</td>
<td>0.80*</td>
</tr>
<tr>
<td>Item #8</td>
<td>0.20</td>
<td>0.20</td>
<td>0.80*</td>
</tr>
<tr>
<td>Item #9</td>
<td>0.20</td>
<td>0.20</td>
<td>0.80*</td>
</tr>
</tbody>
</table>

*Item-scale correlation, corrected for overlap.*
### Item-scale correlation matrix

<table>
<thead>
<tr>
<th></th>
<th>Depress</th>
<th>Anxiety</th>
<th>Anger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item #1</td>
<td>0.50*</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Item #2</td>
<td>0.50*</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Item #3</td>
<td>0.50*</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Item #4</td>
<td>0.50</td>
<td>0.50*</td>
<td>0.50</td>
</tr>
<tr>
<td>Item #5</td>
<td>0.50</td>
<td>0.50*</td>
<td>0.50</td>
</tr>
<tr>
<td>Item #6</td>
<td>0.50</td>
<td>0.50*</td>
<td>0.50</td>
</tr>
<tr>
<td>Item #7</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50*</td>
</tr>
<tr>
<td>Item #8</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50*</td>
</tr>
<tr>
<td>Item #9</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50*</td>
</tr>
</tbody>
</table>

*Item-scale correlation, corrected for overlap.
Validity

• Content validity
  – Patients and/or experts judge the items to be representing the intended concept adequately

• Construct validity
  – Extent to which associations with other variables are consistent with prior hypotheses
Self-Reports of Physical Health Predict Five-Year Mortality

SF-36 Physical Health Component Score (PCS)—T score

% Dead

<table>
<thead>
<tr>
<th>SF-36 PCS Range</th>
<th>Dead (n)</th>
<th>Increase in % Dead</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;35</td>
<td>17 (n=676)</td>
<td>17%</td>
</tr>
<tr>
<td>35-44</td>
<td>6 (n=754)</td>
<td>6%</td>
</tr>
<tr>
<td>45-54</td>
<td>5 (n=1181)</td>
<td>5%</td>
</tr>
<tr>
<td>&gt;55</td>
<td>2 (n=609)</td>
<td>2%</td>
</tr>
</tbody>
</table>

Mortality Prediction with a Single General Self-Rated Health Question

## Evaluating Construct Validity

<table>
<thead>
<tr>
<th>Scale</th>
<th>Age</th>
<th>Obesity</th>
<th>ESRD</th>
<th>Nursing Home Resident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Functioning</td>
<td>Medium (-)</td>
<td>Small (-)</td>
<td>Large (-)</td>
<td>Large (-)</td>
</tr>
<tr>
<td>Depressive Symptoms</td>
<td>?</td>
<td>Small (+)</td>
<td>?</td>
<td>Small (+)</td>
</tr>
</tbody>
</table>

Cohen effect size rules of thumb (d = 0.2, 0.5, and 0.8):
- Small correlation = 0.100
- Medium correlation = 0.243
- Large correlation = 0.371

\[
r = \frac{d}{\sqrt{(d^2 + 4)^{.5}}} = \frac{0.8}{\sqrt{(0.8^2 + 4)^{.5}}} = \frac{0.8}{\sqrt{(0.64 + 4)^{.5}}} = \frac{0.8}{\sqrt{4.64^{.5}}} = \frac{0.8}{2.154} = 0.371
\]

*Beware: r’s of 0.10, 0.30 and 0.50 are often cited as small, medium, and large.*
Responsiveness to Change

• Valid measures should be responsive to interventions that change the thing being measured.

• Compare change on measure to change indicated on external indicator of change ("anchor")
Listed below are a few statements about your relationships with others. How much is each statement TRUE or FALSE for you?

- I am always courteous even to people who are disagreeable.
- There have been occasions when I took advantage of someone.
- I sometimes try to get even rather than forgive and forget.
- I sometimes feel resentful when I don’t get my way.
- No matter who I’m talking to, I’m always a good listener.

Definitely True/Mostly True/Don’t Know/Mostly False/Definitely False
Break #2
Use of Patient-Reported Outcome Measures in Research

11:00-11:50am
Specific Aims

Among Medicare managed care beneficiaries …

- 1) Do the associations of different types of cancer and (non-cancer) chronic conditions with health-related quality of life vary among Medicare managed care beneficiaries?

- 2) Do the associations between cancer and health-related quality of life vary by stage of disease?

• Surveillance, Epidemiology and End Results (SEER) program of cancer registries that collect standardized clinical and demographic information for persons with newly diagnosed (incident) cancer in specific geographical areas.

• Began in 1973 and covers ~ 26% of U.S. pop.
  - [http://seer.cancer.gov/registries/list.html](http://seer.cancer.gov/registries/list.html)
  - California, Connecticut, Hawaii, Iowa, Kentucky, Louisiana, New Mexico, New Jersey, Utah
  - Atlanta, Detroit, rural Georgia, Seattle-Puget Sound metropolitan areas
• Medicare Health Outcomes Survey (MHOS)
  - 95-item survey administered to 1,000 randomly selected beneficiaries (including institutionalized and disabled) in Medicare managed care plans
  - Baseline and follow-up survey (2 years later).
  - 63-72% response rates for baseline surveys

Sample (n = 126,366)

- 55% female
- 79% non-Hispanic white, 7% Hispanic, 5% Black, 5% Asian
- 60% married
- 58% high school graduate or less
- 51% < $30,000 income
Dependent Variable = SF-6D

- SF-36 health survey, version 1

- 11 of 36 questions representing 6 of 8 domains
  - Physical functioning
  - Role limitations
  - Social function
  - Pain
  - Emotional well-being
  - Energy/fatigue

- Standard gamble elicitation of preferences from a population sample in the UK.

- Scores for those alive range from 0.30 to 1.00 (dead = 0.00).
10 Cancer Conditions (n = 22,740; 18%)

- Prostate cancer (n = 5,593; 4%)
- Female breast Cancer (n = 4,311; 3%)
- Colorectal cancer (n = 3,012; 2%)
- Non-small cell lung cancer (n = 1,792; 1%)
- Bladder cancer (n = 1,299; 1%)
- Melanoma (n = 1,135; 1%)
- Endometrial cancer (n = 902; 1%)
- Non-Hodgkin’s lymphoma (n = 668; 1%)
- Kidney cancer (n = 488; 0.4%)
- Other cancer (n = 3,540; 3%)

Note: Those with more than one cancer diagnosis are excluded.
Historic Stage of Disease (time of diagnosis)

• Localized
  – 2045 breast, 2652 prostate, 1481 colorectal, 466 lung

• Distant (metastatic)
  – 26 breast, 61 prostate, 48 colorectal, 47 lung

• Unstaged
  – 347 breast, 633 prostate, 203 colorectal, 65 lung
13 Non-cancer Conditions
(mean number = 2.44)

- Hypertension  
  n = 66,968  (53%)
- Arthritis of the hip  
  n = 44,524  (35%)
- Arthritis of the hand  
  n = 40,402  (32%)
- Sciatica  
  n = 26,878  (21%)
- Other heart disease  
  n = 25,455  (20%)
- Diabetes  
  n = 20,089  (16%)
- Angina/coronary artery disease  
  n = 18,017  (14%)
- Chronic obstructive pulmonary disease  
  n = 15,445  (12%)
- Depressed in the last year  
  n = 14,815  (12%)
- Myocardial infarction/heart attack  
  n = 11,982  (9%)
- Stroke  
  n = 9,479  (8%)
- Congestive heart failure  
  n = 7,893  (6%)
- Inflammatory bowel disease  
  n = 5,882  (5%)

Has a doctor ever told you that you had: …
In the past year, have you felt depressed or sad much of the time?
Demographic & Administration Variables

- Age (continuous)
- Education (8th grade or less; some high school; high school graduate; some college; 4 year college grad; > 4 year college)
- Gender (male; female)
- Income (<10k, 10-19999, 20-29999, 30-39999, 40-49999, 50-79999, 80k and above, don’t know or missing)
- Race/ethnicity (Hispanic, non-Hispanic white, black, Asian, American Indian, other race, missing)
- Marital status (married, widowed, divorced/separated/never married)

- Proxy completed survey (11%)
- Mode of administration (88% mail vs. 12% phone)
Results

• Adjusted R-squared of 39% for 43 dfs
• Intercept = 0.81
  – No chronic condition, average education and age, divorced/separated/never married, white, don’t know/missing income, phone mode)
  – SD = 0.14
• Only 2 of 23 conditions had non-significant associations (melanoma, endometrial cancer)
HRQOL in SEER-Medicare Health Outcomes Study (n = 126,366)

Controlling for age, gender, race/ethnicity, education, income, and marital status.
Distant stage of cancer associated with 0.05-0.10 lower SF-6D Score

Figure 1. Distant Stage of Disease Associated with Worse SF-6D Scores (Sample sizes for local/regional, distant, and unstaged: Breast (2045, 26, 347); Prostate (2652, 61 and 633), Colorectal (1481, 48 and 203), and Lung (466, 47 and 65).
Physical Functioning and Emotional Well-Being at Baseline for 54 Patients at UCLA-Center for East West Medicine

MS = multiple sclerosis; ESRD = end-stage renal disease; GERD = gastroesophageal reflux disease.
Significant Improvement in all but 1 of SF-36 Scales (Change is in T-score metric)

<table>
<thead>
<tr>
<th></th>
<th>Change</th>
<th>t-test</th>
<th>prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF-10</td>
<td>1.7</td>
<td>2.38</td>
<td>.0208</td>
</tr>
<tr>
<td>RP-4</td>
<td>4.1</td>
<td>3.81</td>
<td>.0004</td>
</tr>
<tr>
<td>BP-2</td>
<td>3.6</td>
<td>2.59</td>
<td>.0125</td>
</tr>
<tr>
<td>GH-5</td>
<td>2.4</td>
<td>2.86</td>
<td>.0061</td>
</tr>
<tr>
<td>EN-4</td>
<td>5.1</td>
<td>4.33</td>
<td>.0001</td>
</tr>
<tr>
<td>SF-2</td>
<td>4.7</td>
<td>3.51</td>
<td>.0009</td>
</tr>
<tr>
<td>RE-3</td>
<td>1.5</td>
<td>0.96</td>
<td>.3400</td>
</tr>
<tr>
<td>EWB-5</td>
<td>4.3</td>
<td>3.20</td>
<td>.0023</td>
</tr>
<tr>
<td>PCS</td>
<td>2.8</td>
<td>3.23</td>
<td>.0021</td>
</tr>
<tr>
<td>MCS</td>
<td>3.9</td>
<td>2.82</td>
<td>.0067</td>
</tr>
</tbody>
</table>
Effect Size

$(\text{Follow-up} - \text{Baseline})/ \text{SD}_{\text{baseline}}$

*Cohen’s Rule of Thumb:*

- ✓ $ES = 0.20$   Small
- ✓ $ES = 0.50$   Medium
- ✓ $ES = 0.80$   Large
Effect Sizes for Changes in SF-36 Scores

Effect Size

PFI = Physical Functioning; Role-P = Role-Physical; Pain = Bodily Pain; Gen H=General Health; Energy = Energy/Fatigue; Social = Social Functioning; Role-E = Role-Emotional; EWB = Emotional Well-being; PCS = Physical Component Summary; MCS = Mental Component Summary.
Break #3
Use of Patient-Reported Outcome Measures in Clinical Practice

12:00-12:30 pm
Defining a Responder: Reliable Change Index (RCI)

\[
\frac{X_2 - X_1}{\sqrt{2} \times (SEM)}
\]

\[
SEM = SD_{bl} \times \sqrt{1 - r_{xx}}
\]

*Note: SD_{bl} = standard deviation at baseline \newline r_{xx} = reliability*
Amount of Change Needed for Significant Individual Change

Effect Size

PFI = Physical Functioning; Role-P = Role-Physical; Pain = Bodily Pain; Gen H=General Health; Energy = Energy/Fatigue; Social = Social Functioning; Role-E = Role-Emotional; EWB = Emotional Well-being; PCS = Physical Component Summary; MCS =Mental Component Summary.
7-31% of People in Sample Improve Significantly

<table>
<thead>
<tr>
<th></th>
<th>% Improving</th>
<th>% Declining</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF-10</td>
<td>13%</td>
<td>2%</td>
<td>+ 11%</td>
</tr>
<tr>
<td>RP-4</td>
<td>31%</td>
<td>2%</td>
<td>+ 29%</td>
</tr>
<tr>
<td>BP-2</td>
<td>22%</td>
<td>7%</td>
<td>+ 15%</td>
</tr>
<tr>
<td>GH-5</td>
<td>7%</td>
<td>0%</td>
<td>+ 7%</td>
</tr>
<tr>
<td>EN-4</td>
<td>9%</td>
<td>2%</td>
<td>+ 7%</td>
</tr>
<tr>
<td>SF-2</td>
<td>17%</td>
<td>4%</td>
<td>+ 13%</td>
</tr>
<tr>
<td>RE-3</td>
<td>15%</td>
<td>15%</td>
<td>0%</td>
</tr>
<tr>
<td>EWB-5</td>
<td>19%</td>
<td>4%</td>
<td>+ 15%</td>
</tr>
<tr>
<td>PCS</td>
<td>24%</td>
<td>7%</td>
<td>+ 17%</td>
</tr>
<tr>
<td>MCS</td>
<td>22%</td>
<td>11%</td>
<td>+ 11%</td>
</tr>
</tbody>
</table>
Item Responses and Trait Levels

www.nihpromis.org
Computer Adaptive Testing (CAT)
Reliability Target for Use of Measures with Individuals

- Reliability ranges from 0-1
  - 0.90 or above is goal
    - SEM = SD (1- reliability)\(^{1/2}\)
    - 95% CI = true score +/- 1.96 x SEM
      - if true z-score = 0, then CI: -.62 to +.62
      - Width of CI is 1.24 z-score units

- Reliability = 0.90 when SE = 3.2
  - T-scores (mean = 50, SD = 10)
    \[ T = 50 + (z \times 10) \]
  - Reliability = 1 - (SE/10)^2
Reliability and SEM

• For z-scores (mean = 0 and SD = 1):
  - Reliability = 1 - SE²
  - So reliability = 0.90 when SE = 0.32

• For T-scores (mean = 50 and SD = 10):
  - Reliability = 1 - (SE/10)²
  - So reliability = 0.90 when SE = 3.2
In the past 7 days ...

I was grouchy [1st question]

- Never [39]
- Rarely [48]
- Sometimes [56]
- Often [64]
- Always [72]

Estimated Anger = 56.1
SE = 5.7 (rel. = 0.68)
In the past 7 days ...

I felt like I was ready to explode

[2nd question]
- Never
- Rarely
- Sometimes
- Often
- Always

Estimated Anger = 51.9
SE = 4.8 (rel. = 0.77)
In the past 7 days ...

I felt angry [3rd question]
- Never
- Rarely
- Sometimes
- Often
- Always

Estimated Anger = 50.5
SE = 3.9 (rel. = 0.85)
In the past 7 days ...

I felt angrier than I thought I should
[4th question]

- Never
- Rarely
- Sometimes
- Often
- Always

Estimated Anger = 48.8
SE = 3.6 (rel. = 0.87)
In the past 7 days ...

I felt annoyed [5th question]
- Never
- Rarely
- Sometimes
- Often
- Always

Estimated Anger = 50.1
SE = 3.2 (rel. = 0.90)
In the past 7 days ...

I made myself angry about something just by thinking about it. [6th question]

- Never
- Rarely
- Sometimes
- Often
- Always

Estimated Anger = 50.2
SE = 2.8 (rel = 0.92)
PROMIS Physical Functioning vs. "Legacy" Measures

The Lower the SE: The Greater the Information Content

Mean: U.S. General Population

Worse Physical Function Better Physical Function
“Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations”


- HRQOL has rarely been collected in a standardized fashion in routine clinical practice.
- Increased interest in using PROs for individual patient management.
- Research shows that use of PROs:
  - Improves patient-clinician communication
  - May improve outcomes
Thank you

Ron Hays

drhays@ucla.edu (310-794-2294). Powerpoint file available for downloading at: http://gim.med.ucla.edu/FacultyPages/Hays/

Disclosures. This presentation was supported by NIA Grant P30-AG021684. Dr. Hays receives no UCLA CTSI support.