"A Critical Look at Health-Related Quality of Life Measures"

> SGIM Annual Meeting Ron D. Hays May 2, 2003 (12:30-1:30 pm)

#### ssues

- **1. Too many measures**
- 2. Measurement development standards<sup>^</sup>
- 3. Where is social health?
- 4. Are symptoms different from HRQOL?
- 5. Use of measures with challenging populations
- 6. MID/Responsiveness
- 7. Summarizing profile scores

# Too Many Measures Example from HIV Research

SCALE	SF-56	SF-38	SF-36	SF-31	SF-30	SF-20	SF-20*	HOPES**
1. Physical Function/								
Disability	6	6	10	11	6	6	4	8
2. Role Limitations	2	2	7	0	2	2	2	4
3. Pain	2	1	2	0	1	1	2	3
4. Emotional Distress/wb	22	5	5	5	5	5	3	6
5. Cognitive Distress	6	6	0	5	4	0	2	3
6. Social Function	2	1	2	3	1	1	2	18
7. Fatigue	5	4	4	4	4	0	2	1
8. Health Percentions	5	5	5	2	5	5	3	0
9 Sexual Function/Ool	3	0	0	0	0	0	0	13
9. Overall Quality of Life	3	7	0	0	1	0	0	0

SF-56 = Hays, et al.; SF-38 = Bozzette et al.; SF-36 = Ware & Sherbourne; SF-31 = AIDS HAQ= Lubeck & Fries; SF-30 = MOS-HIV = Wu et al.; SF-20 = Stewart et al.; SF-20\* = Bozzette et al.; \*\*HOPES = Schag and Ganz Measurement Development Standards ^

**Focus Groups** 

**Cognitive Interviews** 

**Readability Assessment** 

**Field Testing** 

**CTT/IRT** 

# Cultural Adaptation of Survey Instruments



# **Model of Response Processes**

Comprehension Understanding the questions Retrieval Recalling information Judgment Deciding relevance Response Formulating answers

Tourangeau, 1984

# **Concurrent Think Aloud**

Respondent verbalizes thoughts while going through the survey.

Interviewers encourages subject to think aloud:

- "Tell me what you are thinking"
- "Say more about that"
- **Specific probes** 
  - "How did you decide to chose that answer?"
  - "What does 'downhearted and blue' mean to you?"

# **Respondent Debriefing**

#### **Retrospective approach**

 Interviewer asks respondent about completing the survey after completing the entire survey or a section of the survey.

#### **Interviewer observation**

 Identifies problematic skip patterns or questions in advance of debriefing

# **Evaluating Multi-item Scales**

**Scale Characteristics** 

**Reliability and unidimensionality** 

**Distribution of scores (level on attribute)** 

**Item Characteristics** 

**Item difficulty** 

Item-scale correlation ("discrimination")

#### Dichotomous Items Showing DIF (2-Parameter Model)



# **Recommended URLs for IRT**

http://appliedresearch.cancer.gov/areas/cognitive/immt.pdf

http://work.psych.uiuc.edu/irt/

http://www.ssicentral.com/home.htm

# **Upcoming Conferences**

October 17-19, *Applications of Item Response Theory to Health*. International Conference on Health Policy Research: Methodological Issues in Health Services and Outcomes Research, Chicago

Spring, 2004. NCI sponsored meeting, *Improving the Measurement of Cancer Outcomes through the Applications of Item Response Theory (IRT) Modeling: Exploration of Item Banks and Computer-Adaptive Assessment*. DC.



# Health-Related Quality of Life (HRQOL)

#### What the person says they can DO (functioning)

- Self-care
- Role
- Social

How the person FEELS (well-being)

- Emotional well-being
- Pain
- Energy

# Are Symptoms Different from HRQOL?

- Symptoms/problems (12 items)
- Effects of kidney disease (8 items)
- Burden of kidney disease (4 items)
- Work status (2 items)
- Cognitive function (3 items)
- Quality of social interaction (3 items)
- Sexual function (2 items)
- Sleep (4 items)

# **Measuring Challenging Populations**

#### Readability

Disabled

# Applying Readability Formulas to Surveys

- Survey instruments stripped of response options.
- Three 100-word samples randomly chosen from survey.
- Word (syllables) and sentence (number) counts averaged across three 100-word samples.
- Formulas applied.



Hays, et al. (2000), <u>American Journal of Medicine</u>

# National Institute on Disability and Rehabilitation Research

"having a disability does not mean that a person is ill" www.ncddr.org/rpp/hf/lrp\_ov.html

"To be healthy does not mean to be free of disease; it means that you can function, do what you want to do, and become what you want to become" (Rene Dubos)

47% of Americans Disabled for Assistance Programs Todays (ADAPT) say they "would not want my disability to be cured"

# Attribution to health vs. disability

Does your health now limit you in these activities?

Persons with disability distinguish health from disability

- "I' m disabled but I feel healthy" (von Faber et al., 2001, p. 2696)
- Whether people focus on their specific condition or exclude it is a general issue (e.g., kidney disease)

## Recommendations

Whatever it means to you

Does your health or a disability now limit you in these activities?

Instructions should also indicate whether to include or exclude disability

**Compare original and new version** 

# Andresen et al. (1999)

Face validity of SF-36 suspect in nursing home setting

 9 items refer to activities not generally performed in this setting (e.g., carrying groceries)

• 6 refer to work

Age and Ageing, 28, 562-566, 1999

Assessing physical functioning Does your health now limit you in (if so, how much) ... climbing several flights of stairs climbing one flight of stairs walking more than a mile walking several hundred yards walking one hundred yards **Offensive to those with mobility impairments** Andresen & Meyers (2000, Archives of Physical Medicine and **Rehabilitation**)

# Reject "walk" and "climb"

#### Gently

- I can't walk, but in my wheelchair, I can go a mile
- I can't climb a flight of stairs, but if there's a ramp, I can get from the first floor to the second

Less gently

- Annoyed by repeated questions in sequence
- Hanging up during phone survey

# Mattson-Prince (1997)

Dropped 10 physical functioning items because of perception that they were demeaning to people with SCI

Spinal Cord, 35, 326-331

# **Possible Actions**

Switch order of items from easiest to hardest and skip out

CAT

Substitute "go" for walk and climb

**Administer both versions** 

Uncorrected and corrected physical function

**Quality of Well-Being Scale** 0.940 = In wheelchair & moved oneself 0.923 = In wheelchair & did not move oneself 0.640 = Average in CGA (65-99 year olds) 0.501 = Spinal cord injury Saving life of "healthy" person yields more QALYs

# **Community should weight states**

Ex ante judgments

Health care subsidized by public

# Those with condition should weight states

Ex post judgments

"Desirability of a condition to people who are not in it themselves is only moderately correlated to the experienced well-being of people with the condition" (Nord, 2001, p. 579)

# **Responsiveness to Change**

- HRQOL measures should be responsive to interventions that change HRQOL
- Evaluating responsiveness requires assessing HRQOL relative to an external indicator of change (anchor)

## **Two Essential Elements**

1. External (not HRQOL measure being evaluated) indicator of change (Anchor)

2. Amount of HRQOL change among those determined to have changed on anchor, relative to noise (variance).

### **Responsiveness Indices**

(1) Effect size (ES) = D/SD

(2) Standardized Response Mean (SRM) = D/SD<sup>†</sup>

(3) Guyatt responsiveness statistic (RS) = D/SD<sup>‡</sup>

D = raw score change in "changed" group; SD = baseline SD; SD<sup>†</sup> = SD of D; SD<sup>‡</sup> = SD of D among "unchanged"

# **Kinds of Anchors**

#### Self-report

Clinician or other report

Clinical parameter

Clinical intervention

# Magnitude of HRQOL Change Should Parallel Underlying Change



->Size of Intervention

# Change and Responsiveness in PCS Depends on Treatment



# Change and Responsiveness in MCS Depends on Treatment



# Minimal Clinically Important Difference (MCID)

Smallest difference in score that is worth caring about (important).

- Some differences between groups or over time may be so small in magnitude that they are not important.

Change large enough for a clinician to base treatment decisions upon it.

# **Two Essential Elements**

1. Indicator (not HRQOL measure being evaluated) of "minimal" change (Anchor)

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

# **Example 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

# MID for SF-36 is:

# "typically in the range of 3 to 5 points" (p. 149). {. 09->0.28 ES}

Samsa, G., Edelman, D., Rothman, M. L., Williams, G. R., Lipscomb, J., & Matchar, D. Pharmacoeconomics, <u>15</u>, 141-155: 1999.

Does this mean that 1-2 points on the SF-36 scales are unimportant?

# **Bottom Line**

Identification of MID aids interpretation by providing familiar anchors to unfamiliar units.

Trying to give a single point estimate is too simplistic.

Bounded estimates are necessary given the uncertainty.

# **MID Varies by Anchor**

693 RA clinical trial participants evaluated at baseline and 6-weeks post-treatment.

Five anchors: 1) patient global self-report; 2) physician global report; 3) pain self-report; 4) joint swelling; 5) joint tenderness

Kosinski, M. et al. (2000). Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. <u>Arthritis and Rheumatism</u>, <u>43</u>, 1478-1487.

# Changes in SF-36 Scores Associated with Minimal Change in Anchors

Scale	Self-R	ClinR	Pain	Swell	Tender	Mean
PF	8	8	8	6	8	8
Role-P	21	20	11	13	13	16
Pain	15	12	8	12	7	11
GH	4	2	2	3	1	2
EWB	7	5	5	3	1	4
Role-E	18	12	8	16	11	16
SF	12	9	8	8	10	9
EF	11	10	5	5	8	8
PCS	4	4	3	3	3	4
MCS	5	3	2	3	2	3

# MID Varies by Starting Position and Direction of Change

Same retrospective report of change associated with bigger prospective change for those with more room to change

Among those who said their physical health was somewhat worse, change ranged from –26 points to +3 points for people with high (81-100) versus low (0-20) baseline physical health (Baker et al., 1997, <u>Medical Care</u>).

# Group Average is Different from Individual Change

Average change collapses across individual responses.

Is it reasonable to infer the minimum amount of change that is important for individuals based on a group average?

What if SF-36 scores improved by 4 points for half the people and 0 points for the other half?

# MID Determination Complicated By Cumulative Change Over time



# Note: 4-year decline in PCS among US seniors, 1990-94.

-> 1.5 points per year (0.15 SD)



# Summarizing Profile Measures



• Taft, C., Karlsson, J., & Sullivan, M. (2001). Do SF-36 component score accurately summarize subscale scores? Quality of Life Research, 10, 395-404.

•Ware, J. E., & Kosinski, M. (2001). Interpreting SF-36 summary health measures: A response. <u>Quality of Life</u> <u>Research</u>, <u>10</u>, 405-413.

•Taft, C., Karlsson, J., & Sullivan, M. (2001). Reply to Drs Ware and Kosinski. <u>Quality</u> of Life Research, 10, 415-420.

# SF-36 Physical Health



# SF-36 Mental Health



## MCS = (PF\_Z \* -.22999) + (RP\_Z \* -.12329) + (BP\_Z \* -.09731) + (GH\_Z \* -.01571) + (EF\_Z \* .23534) + (SF\_Z \* .26876) + (RE\_Z \* .43407) + (EW\_Z \* .48581)

PCS = (PF\_Z \* .42402) + (RP\_Z \* .35119) + (BP\_Z \* .31754) + (GH\_Z \* .24954) + (EF\_Z \* .02877) + (SF\_Z \* -.00753) + (RE\_Z \* -.19206) + (EW\_Z \* -.22069)

### SF-36 PCS and MCS

# 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.
Yet SF-36 PCS did not improve.

♦ Simon et al. (Med Care, 1998)



Four scales improve 0.28-0.49 SD, but physical health summary score doesn't change



# n = 194 with Multiple Sclerosis

Scores than general population on  $\otimes$  Emotional well-being ( $\downarrow 0.3 SD$ ) Role—emotional (↓ 0.7 SD)  $\otimes$ Energy ( $\downarrow$ 1.0 SD)  $\otimes$  Social functioning ( $\downarrow 1.0$  SD) ♦Yet SF-36 MCS was only <u>0.2</u> SD lower. 

Nortvedt et al. (Med Care, 2000)



# SF-36 Factor Analysis in Singapore

	English		Chin	ese	United States	
	Physical	Mental	Physical	Mental	Physical	Mental
PF	0.60	0.14	0.75	0.03	0.85	0.12
RP	0.85	0.12	0.78	0.25	0.81	0.27
BP	0.46	0.53	0.53	0.51	0.76	0.28
GH	0.14	0.74	0.32	0.66	0.69	0.37
VT	0.15	0.84	0.16	0.83	0.47	0.64
SF	0.49	0.56	0.48	0.56	0.42	0.67
RE	0.77	0.18	0.62	0.36	0.17	0.78
MH	0.12	0.83	0.10	0.86	0.17	0.87

# **Contact Information and URLs**

Ron D. Hays, UCLA Division of General Internal Medicine and Health Services Research (email: <u>hays@rand.org</u>)

www.rand.org/health/surveys.html

http://gim.med.ucla.edu/kdqol/

http://gim.med.ucla.edu/facultypages/hays

# A Critical Look at Health-Related Quality Of Life Measures

This meet the professor session by Ron D. Hays, Ph.D., was supported in part by the UCLA/DREW **Project EXPORT, National Institutes of Health, National Center on Minority Health & Health** Disparities, (P20-MD00148-01) and the UCLA Center for Health Improvement in Minority Elders / Resource **Centers for Minority Aging Research, National** Institutes of Health, National Institute of Aging, (AG-02-004).

> **Resource Centers for Minority Aging Research**

