Patient-Reported Outcomes of Health Care

Presentation to Visiting Scholars from
China Academy of Chinese Medical Sciences

April 16, 2014
9:00am – 11:00am
UCLA Center for East-West Medicine
1033 Gayley Ave, Suite 111
Los Angeles, CA 90024

Ron D.Hays, Ph.D.
drhays@ucla.edu

U.S. Health Care Issues

- **A**ccess to care
  - ~ 50 million people without health insurance
- **C**osts of care
  - Expenditures ~ $ 2.7 Trillion
- **E**ffectiveness (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 ↓

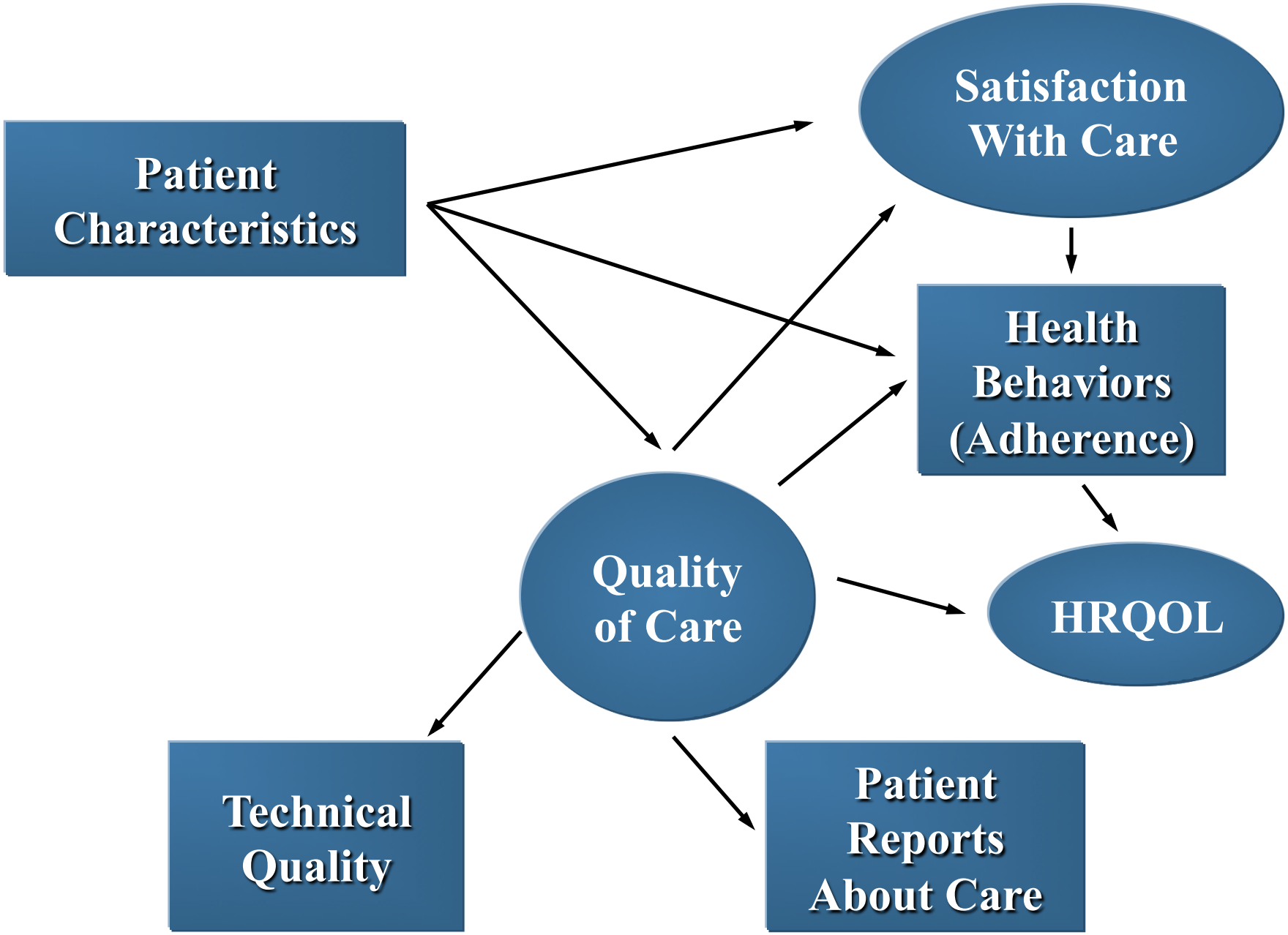
Effectiveness ↑

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)


Patient Characteristics
Satisfaction With Care
Health Behaviors (Adherence)
Quality of Care
HRQOL
Technical Quality
Patient Reports About Care

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

"I'm afraid that your irritable bowel syndrome has progressed. You now have furious and vindictive bowel syndrome."

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

Targeted Multi-Item Scale Burden of Kidney Disease

- 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?

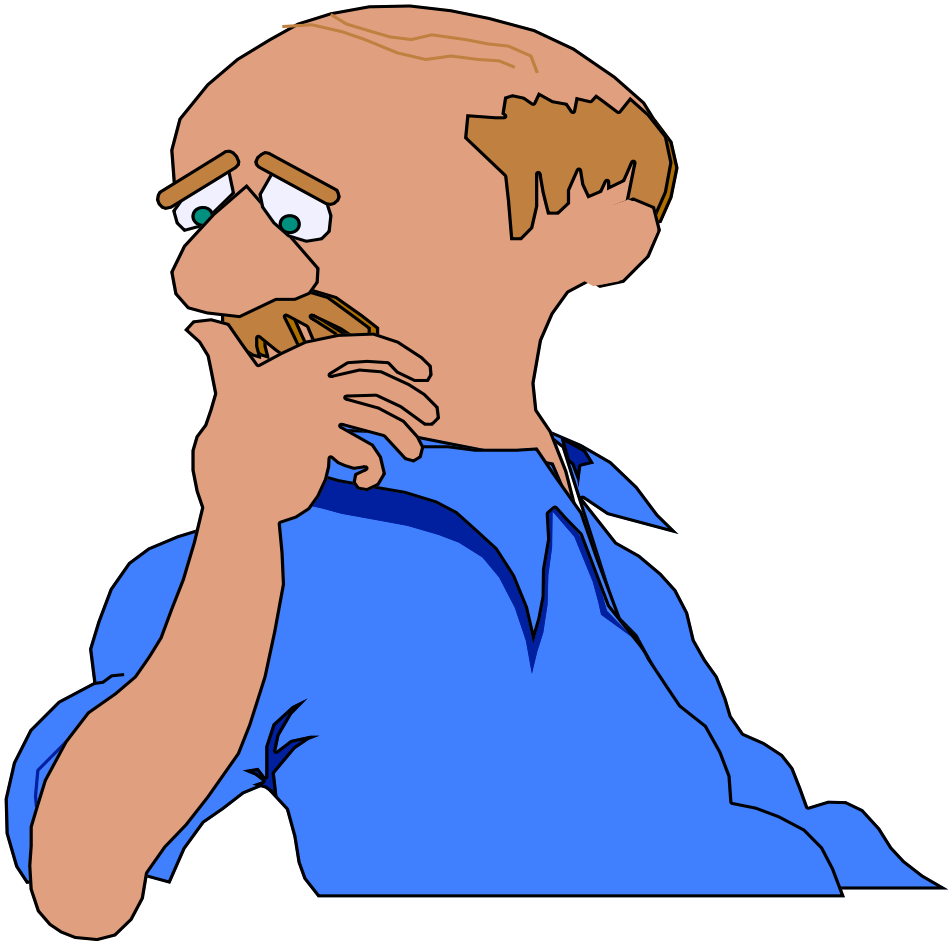
None of the time

A little of the time

Some of the time

Most of the time

All of the time



Generic Scales (SF-36)

- Physical functioning (10 items)
- Role limitations/physical (4 items)
- Emotional well-being (5 items)
- Role limitations/emotional (3 items)
- Social functioning (2 items)
- Pain (2 items)
- Energy/fatigue (4 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

$$X = \frac{\text{(original score - minimum) *100}}{\text{(maximum - minimum)}}$$

$$Y = \text{target mean + (target SD * Zx)}$$

$$Z_X = \frac{(X - \bar{X})}{SD_x}$$

Example of Computing z-score and T-score

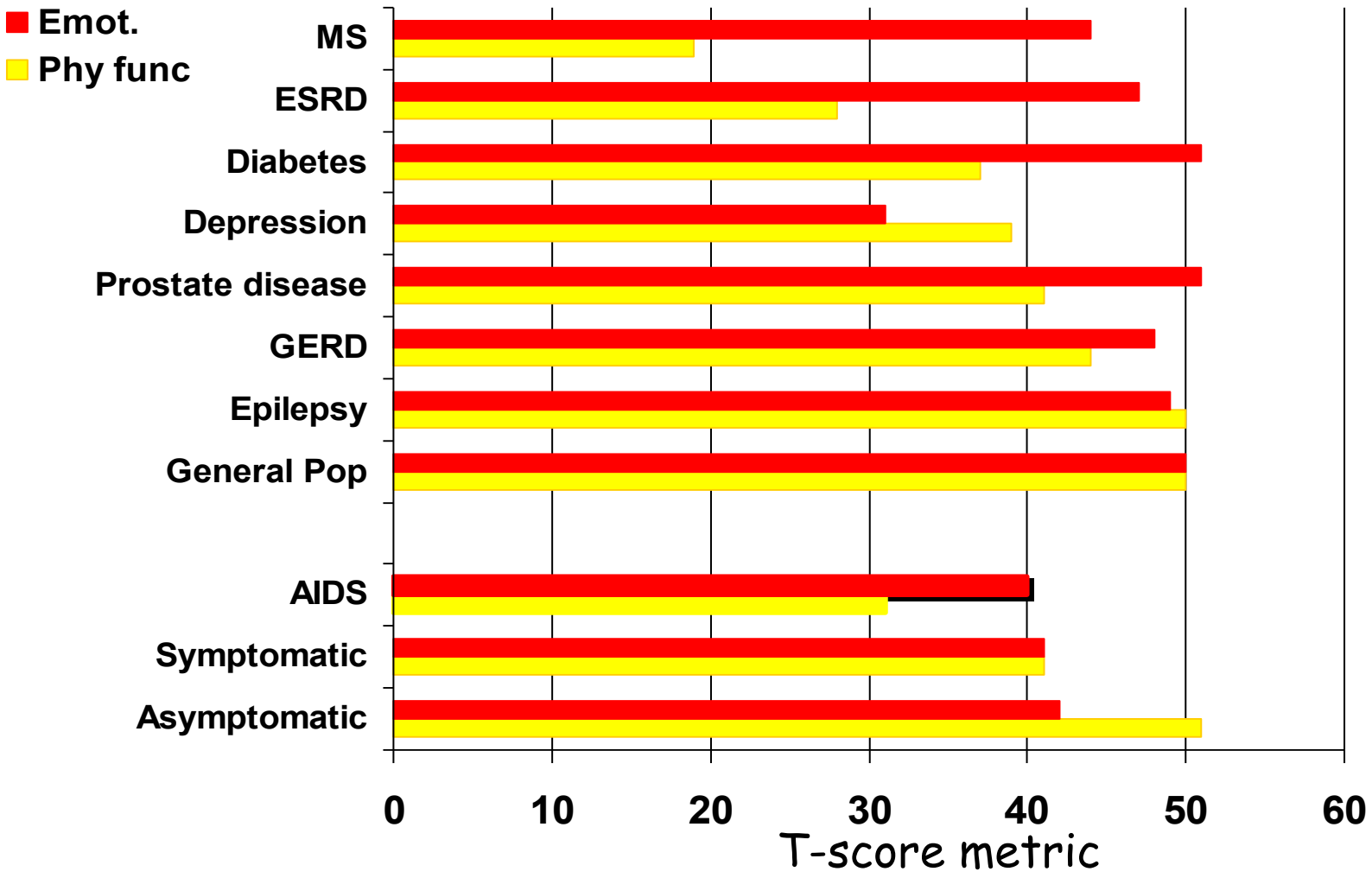
z-score = (score – mean)/SD

T-score = (10 * z-score) + 50

z-score = (100- 36)/31 = 2.06

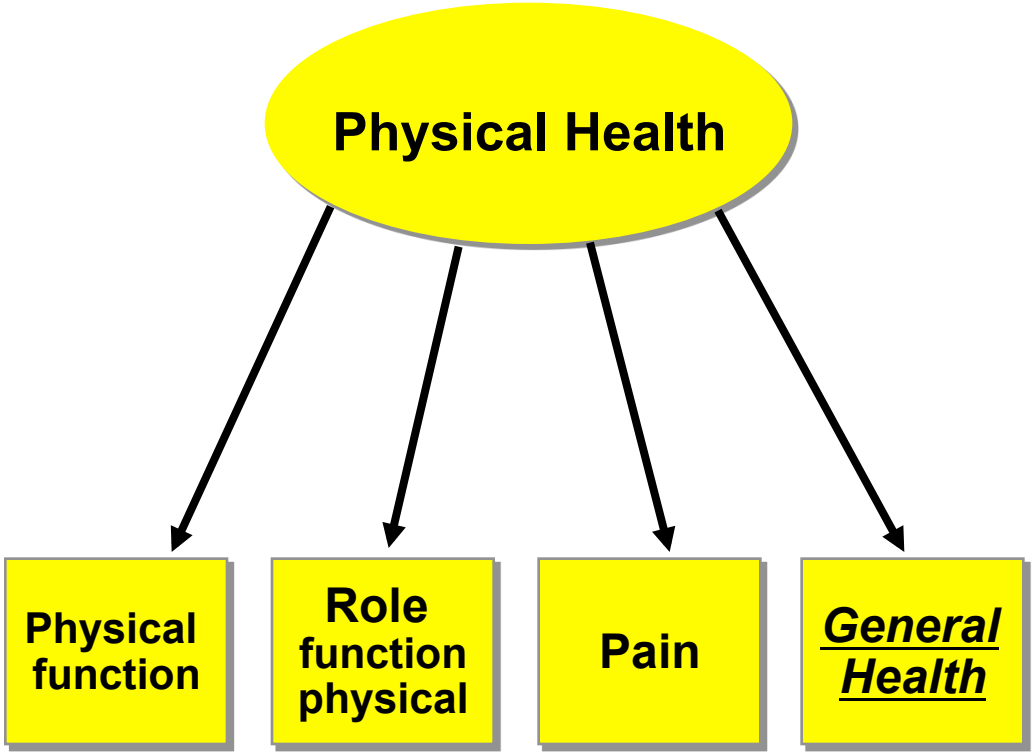
T-score = 71

HRQOL in HIV Compared to other Chronic Illnesses and General Population

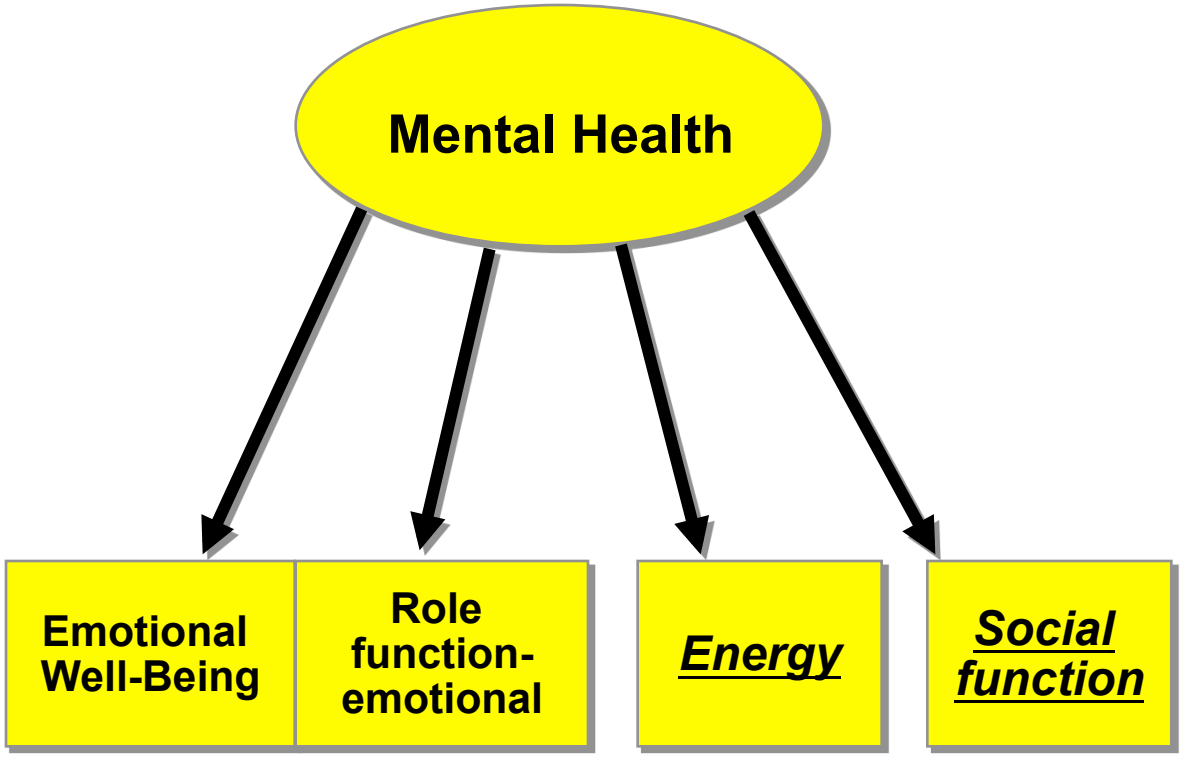


Hays et al. (2000), American Journal of Medicine

Physical Health



Mental Health



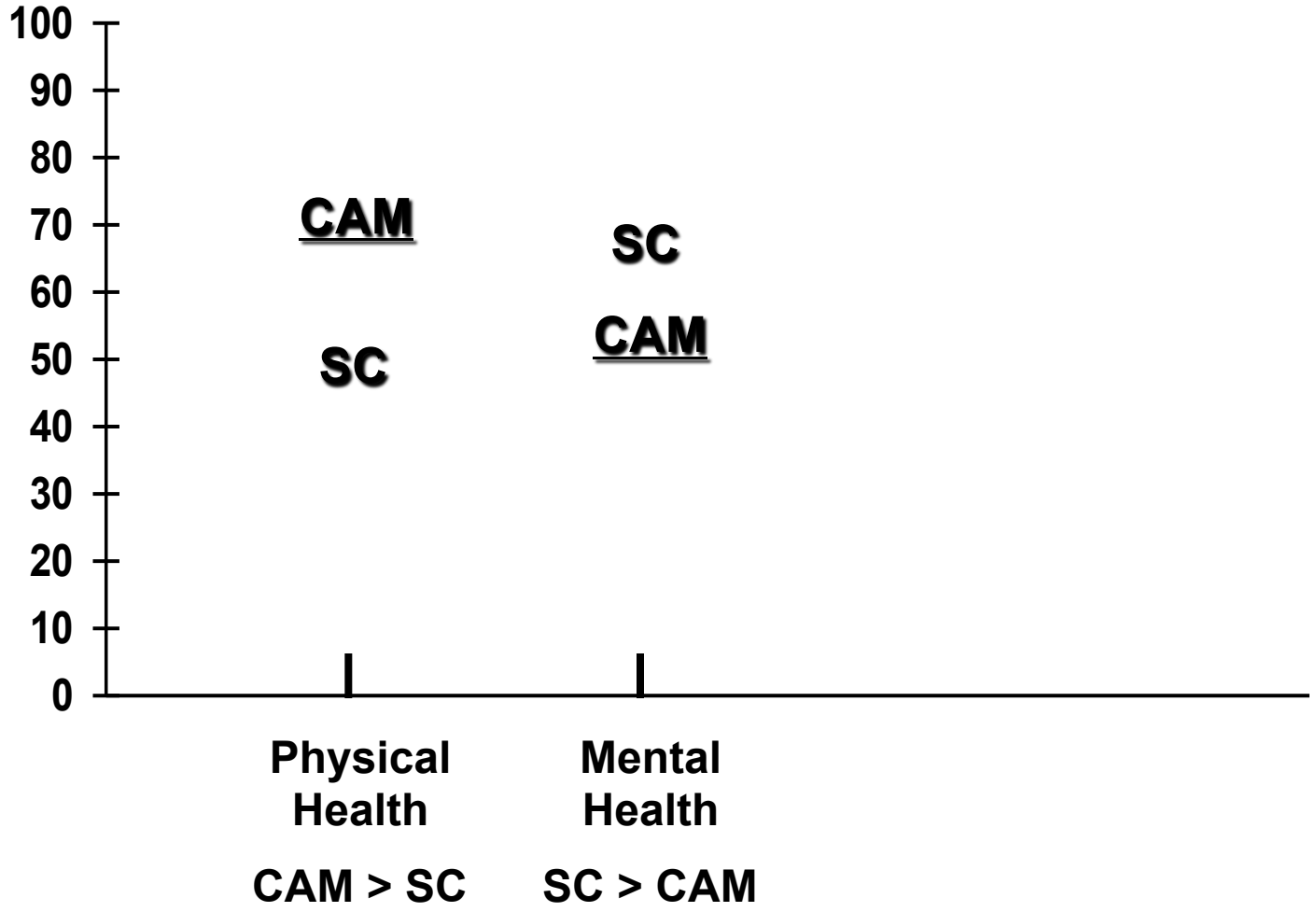
SF-36 PCS and MCS

PCS_z = (PF_Z * 0.42) + (RP_Z * 0.35) +
(BP_Z * 0.32) + (GH_Z * 0.25) +
(EF_Z * 0.03) + (SF_Z * -.01) +
(RE_Z * -.19) + (EW_Z * -.22)

MCS_z = (PF_Z * -.23) + (RP_Z * -.12) +
(BP_Z * -.10) + (GH_Z * -.02) +
(EF_Z * 0.24) + (SF_Z * 0.27) +
(RE_Z * 0.43) + (EW_Z * 0.49)

PCS = (PCS_z*10) + 50
MCS = (MCS_z*10) + 50

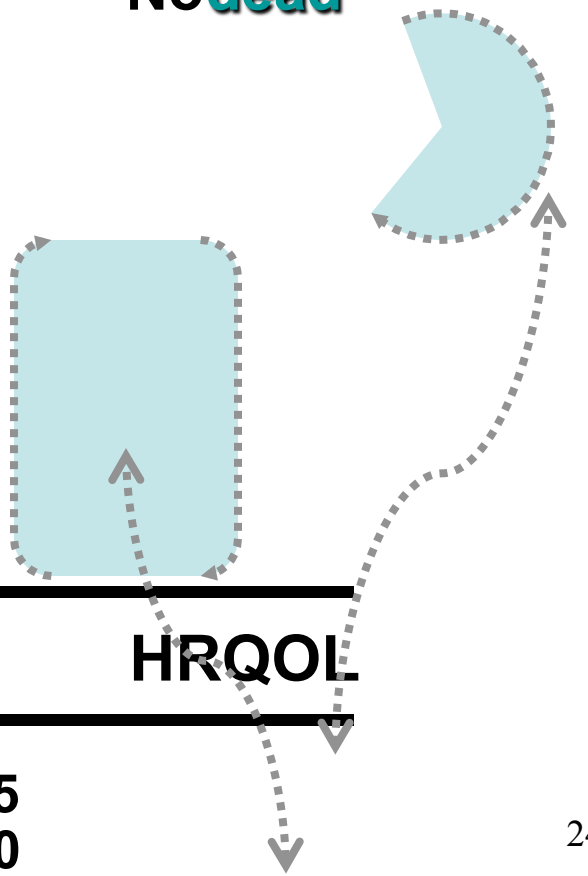
Is CAM Better than Standard Care (SC)?



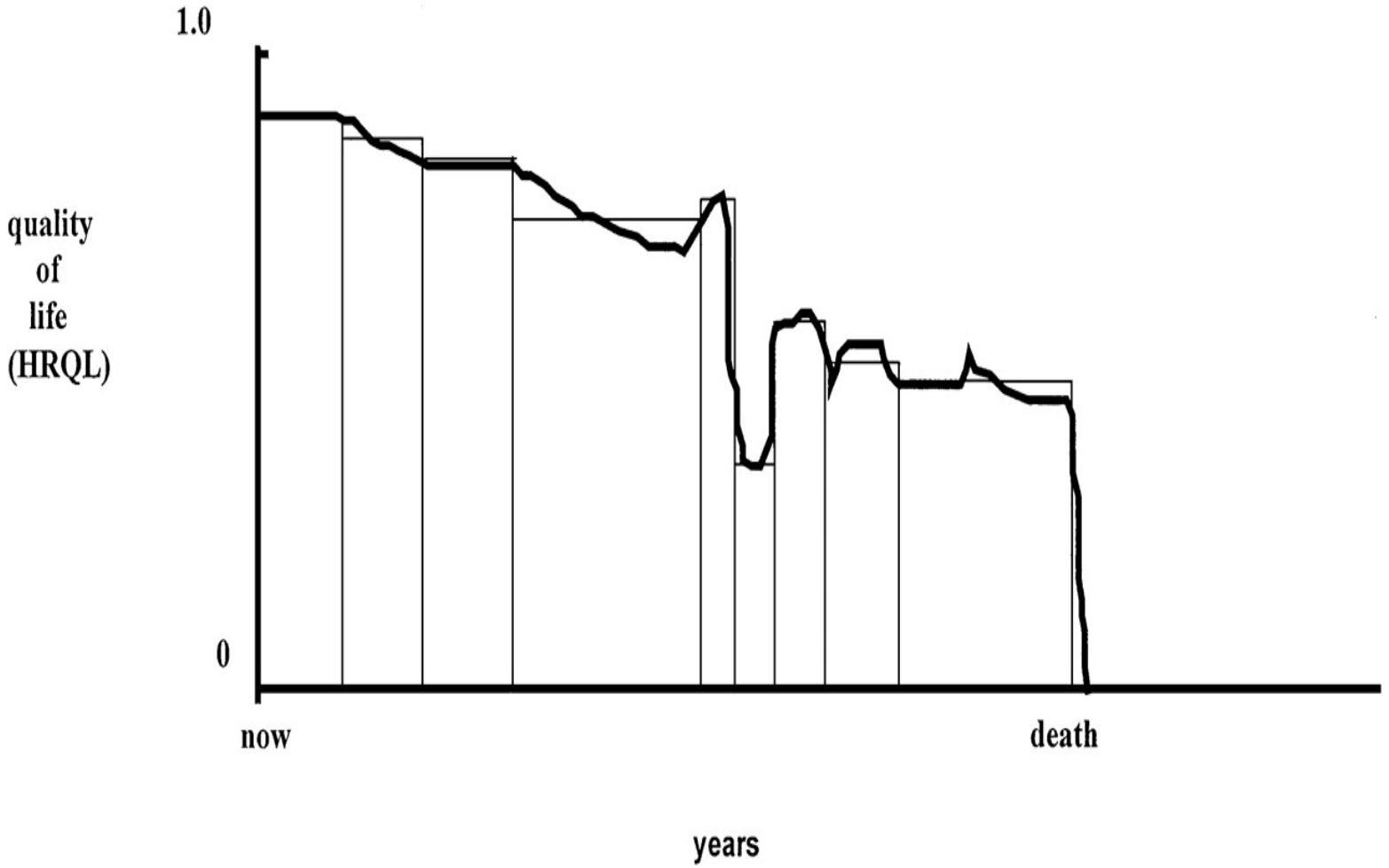
Is Acupuncture Related to Worse HRQOL?

Subject	Acupuncture		HRQOL
	(0-100)		
		1	No dead
		2	No dead
		3 No 50	
		4 No 75	
		5 No 100	
		6 Yes 0	
		7 Yes 25	
		8 Yes 50	
		9 Yes 75	
		10 Yes 100	

Group	n	HRQOL
No Acupuncture	375	
Yes Acupuncture	550	



Quality of Life for Individual Over Time



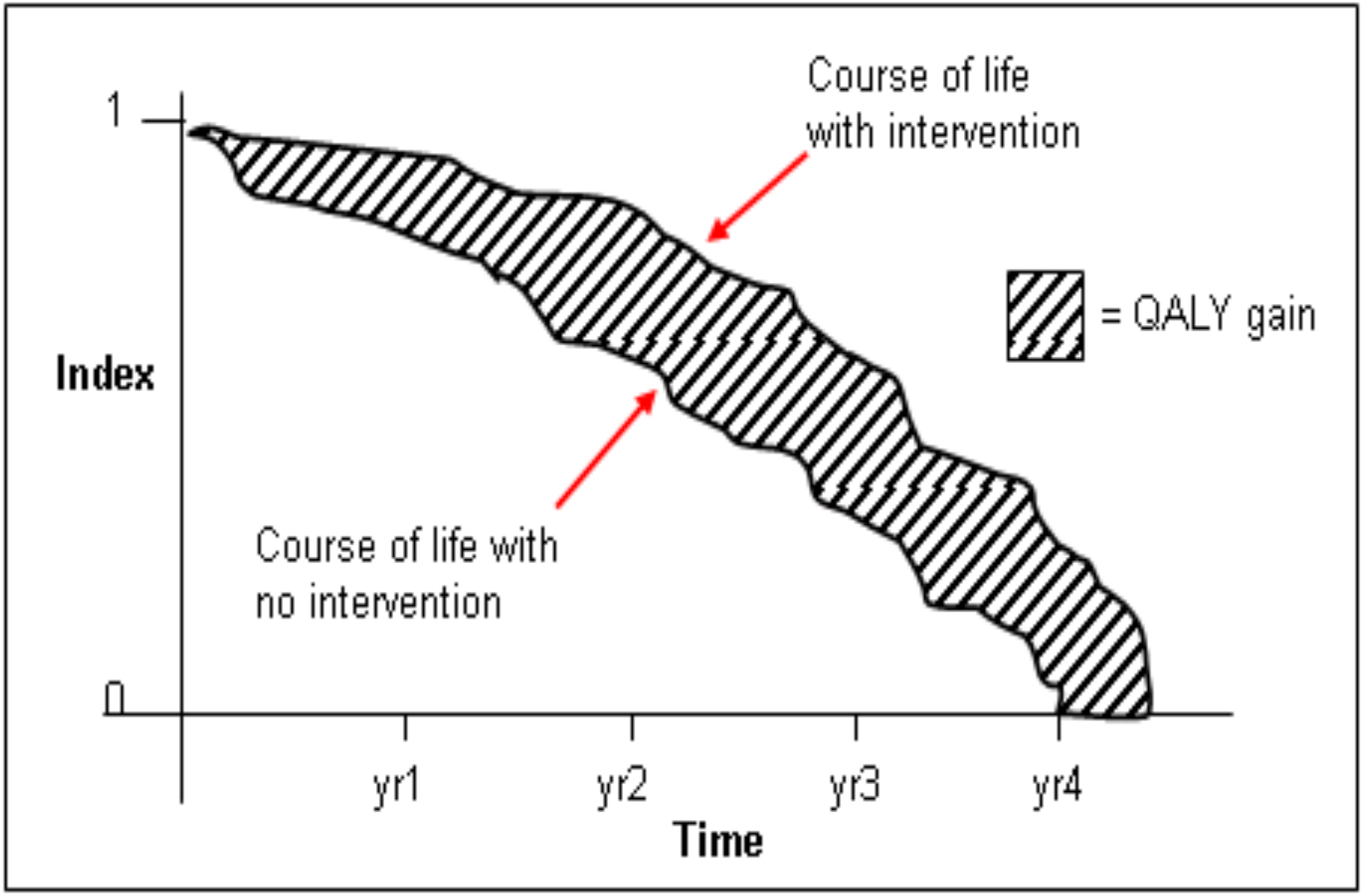
Ultimate Use of HRQOL Measures--Helping to Ensure Access to Cost-Effective Care

Cost ↓

— — — — — — — — —

Effectiveness (“Utility”) ↑

http://www.ukmi.nhs.uk/Research/pharma_res.asp

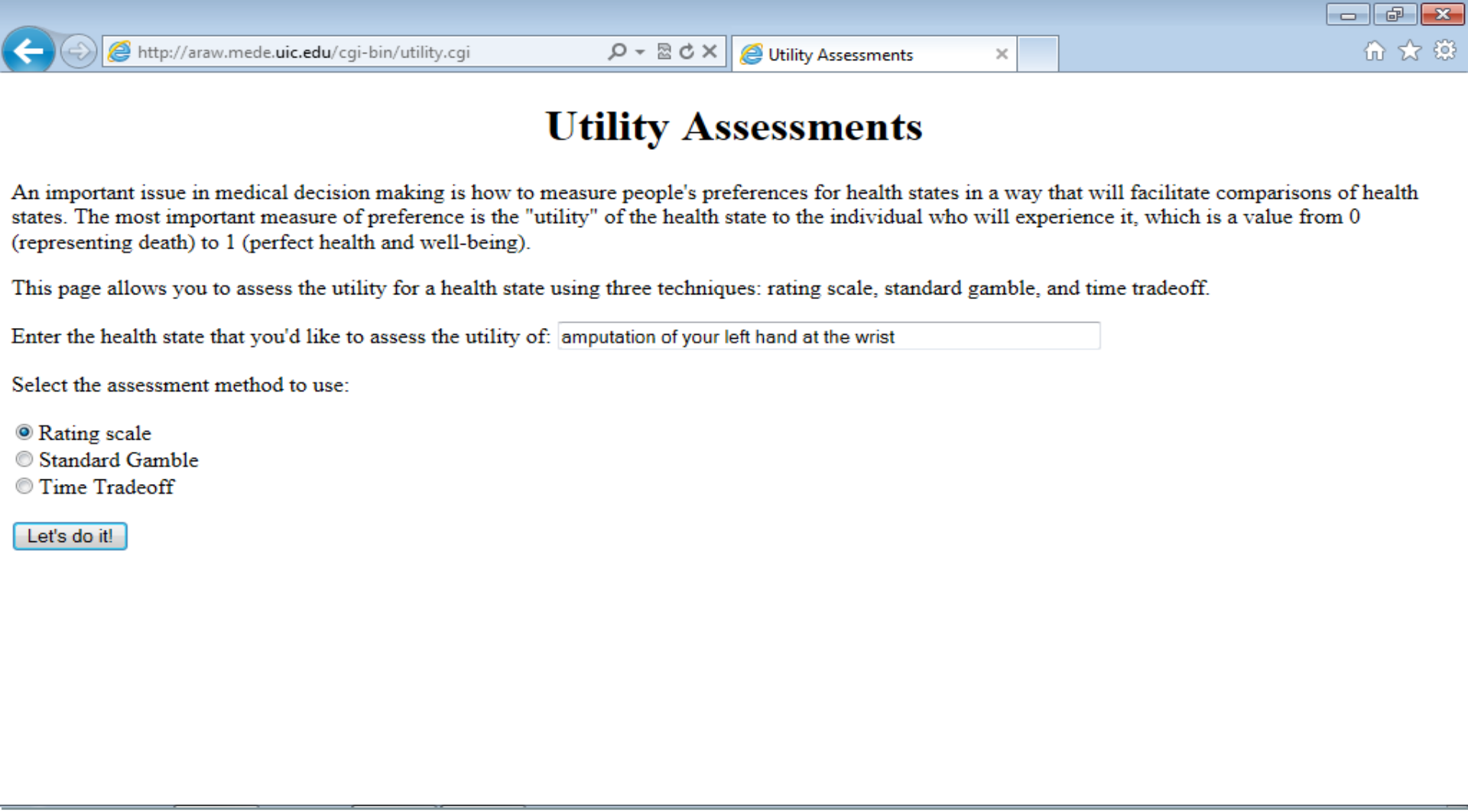


"QALYs: The Basics"

Milton Weinstein, George Torrance, Alistair McGuire

(Value in Health, 2009, vol. 12 Supplement 1)

- 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.



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

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

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

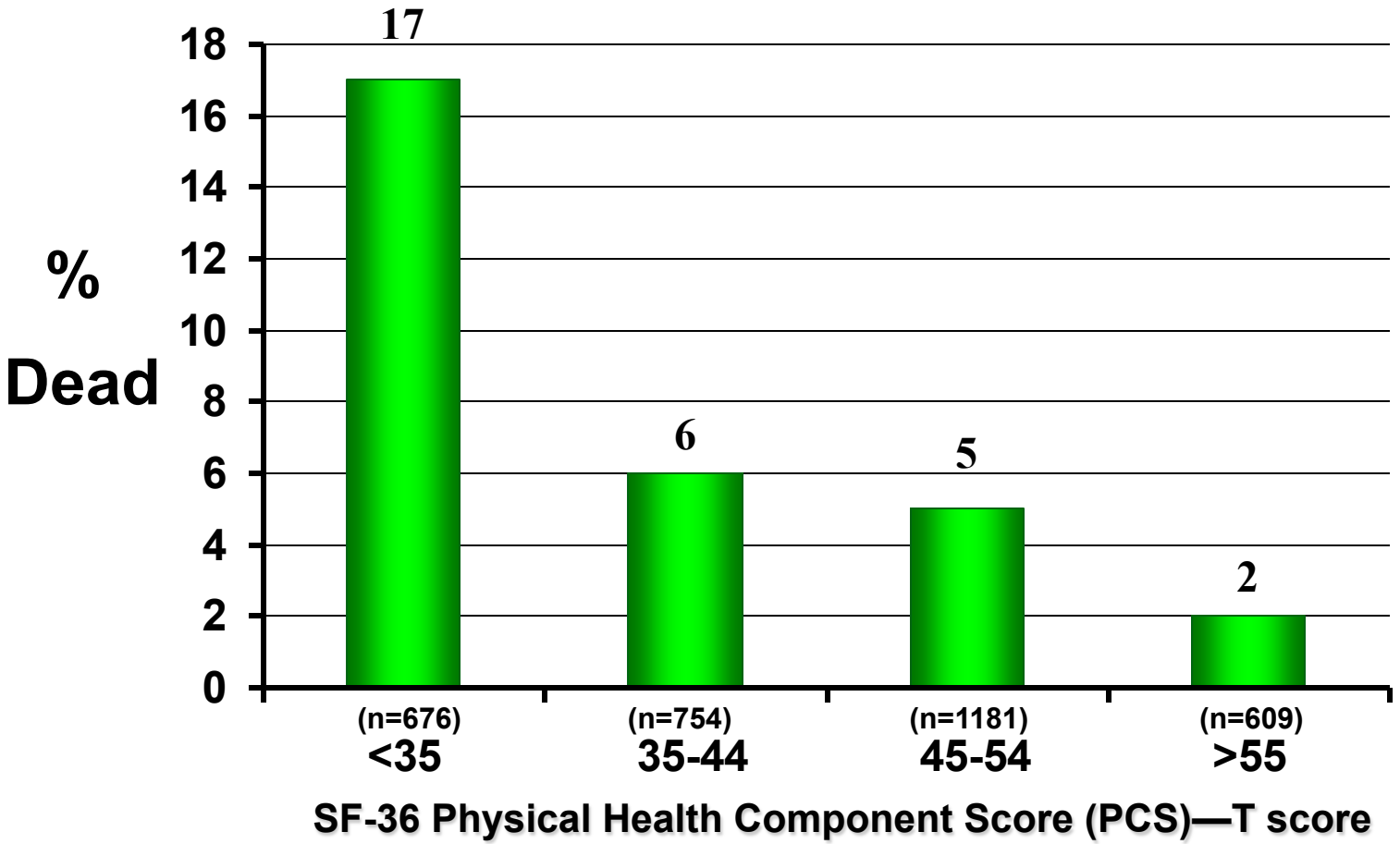
Range of reliability estimates

0.80-0.90 for blood pressure

0.70-0.90 for multi-item self-report scales

Hahn, E. A., Cella, D., et al. (2007). Precision of health-related quality-of-life data compared with other clinical measures. Mayo Clin Proceedings, 82 (10), 1244-1254.

Self-Reports of Physical Health Predict Five-Year Mortality

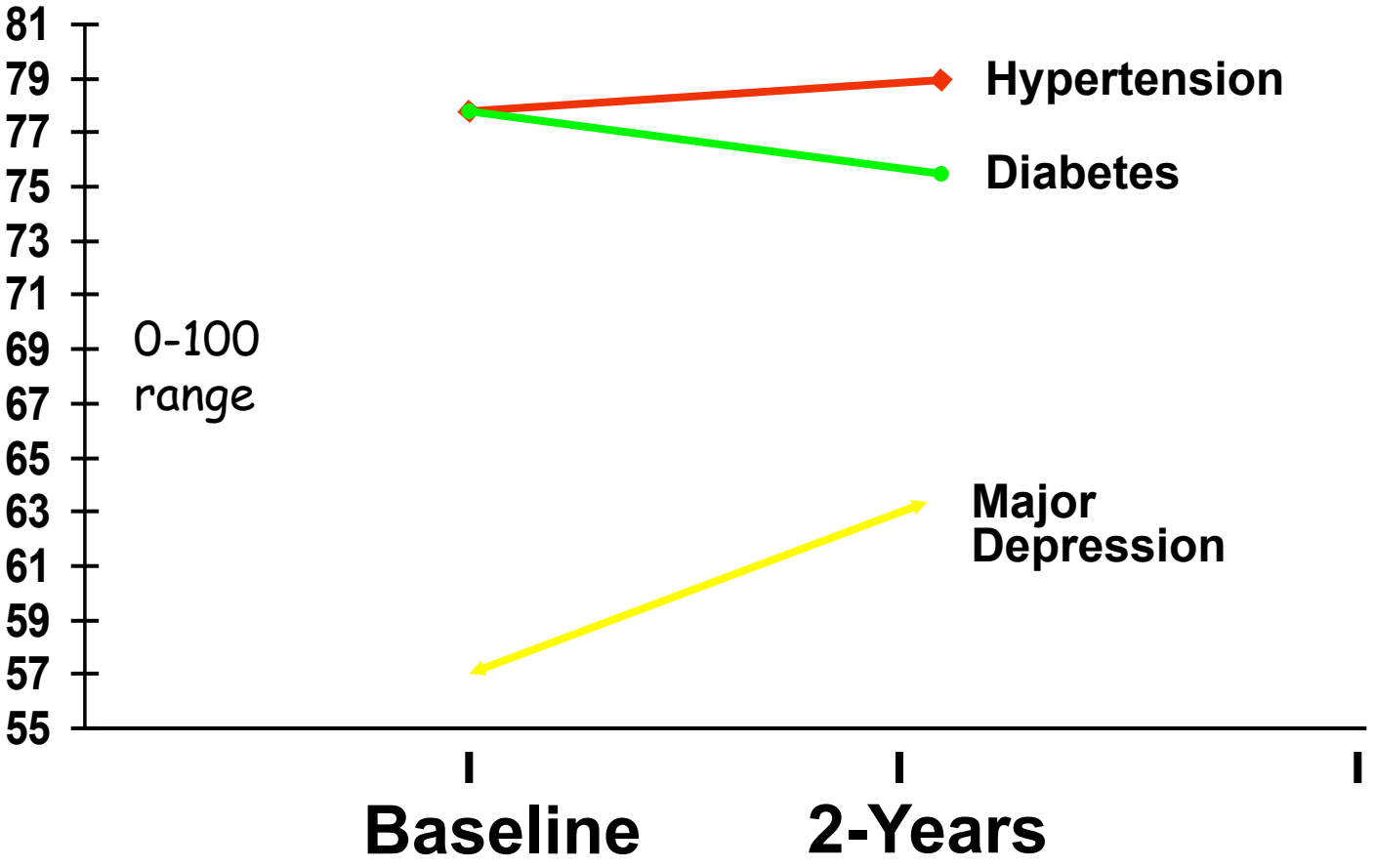


Ware et al. (1994). SF-36 Physical and Mental Health Summary Scales: A User's Manual.

Mortality Prediction with a Single General Self-Rated Health Question

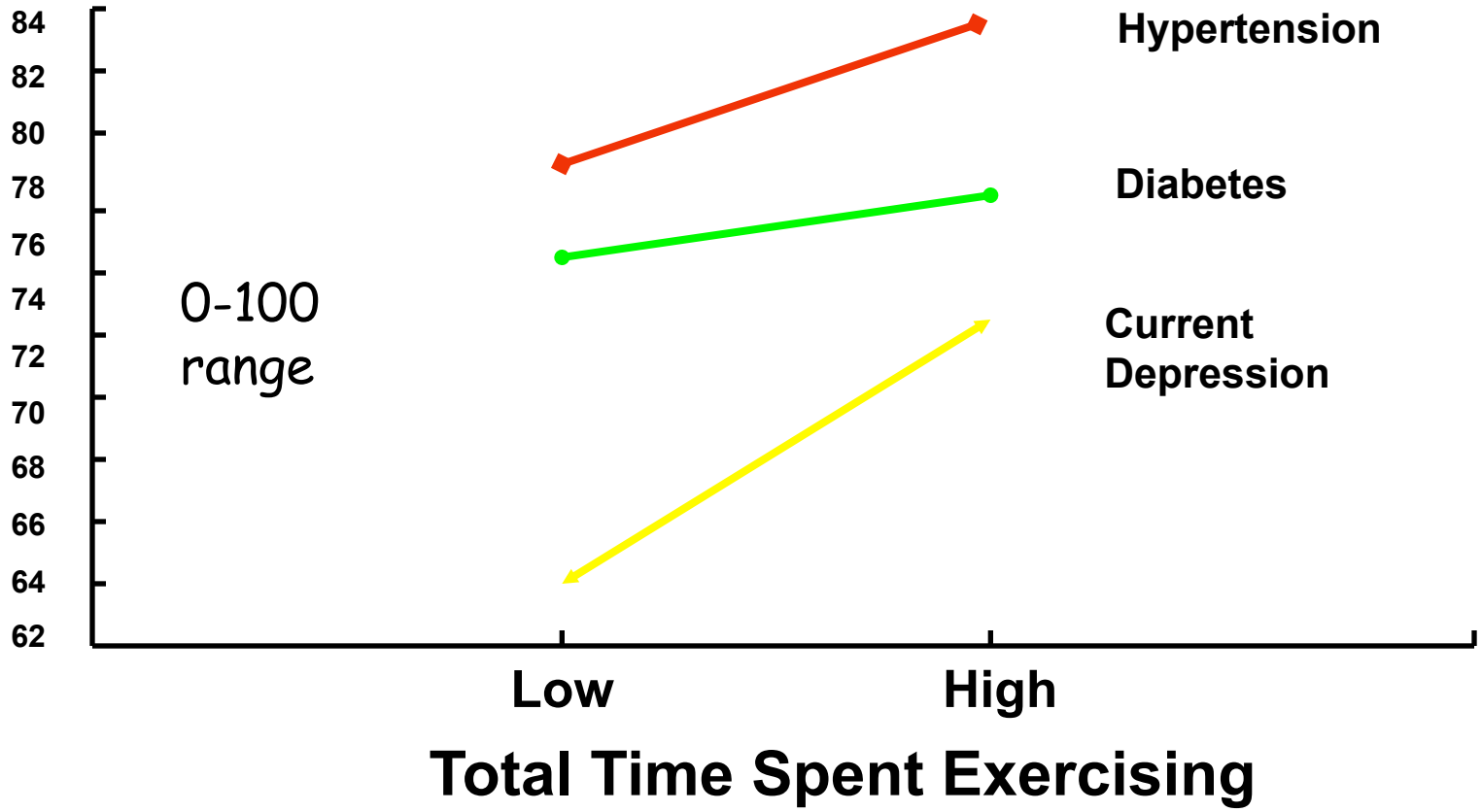
DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2005). Mortality prediction with a single general self-rated health question: A meta-analysis. JGIM, 20, 267-275.

Course of Emotional Well-being Over 2-years for Patients in the MOS General Medical Sector



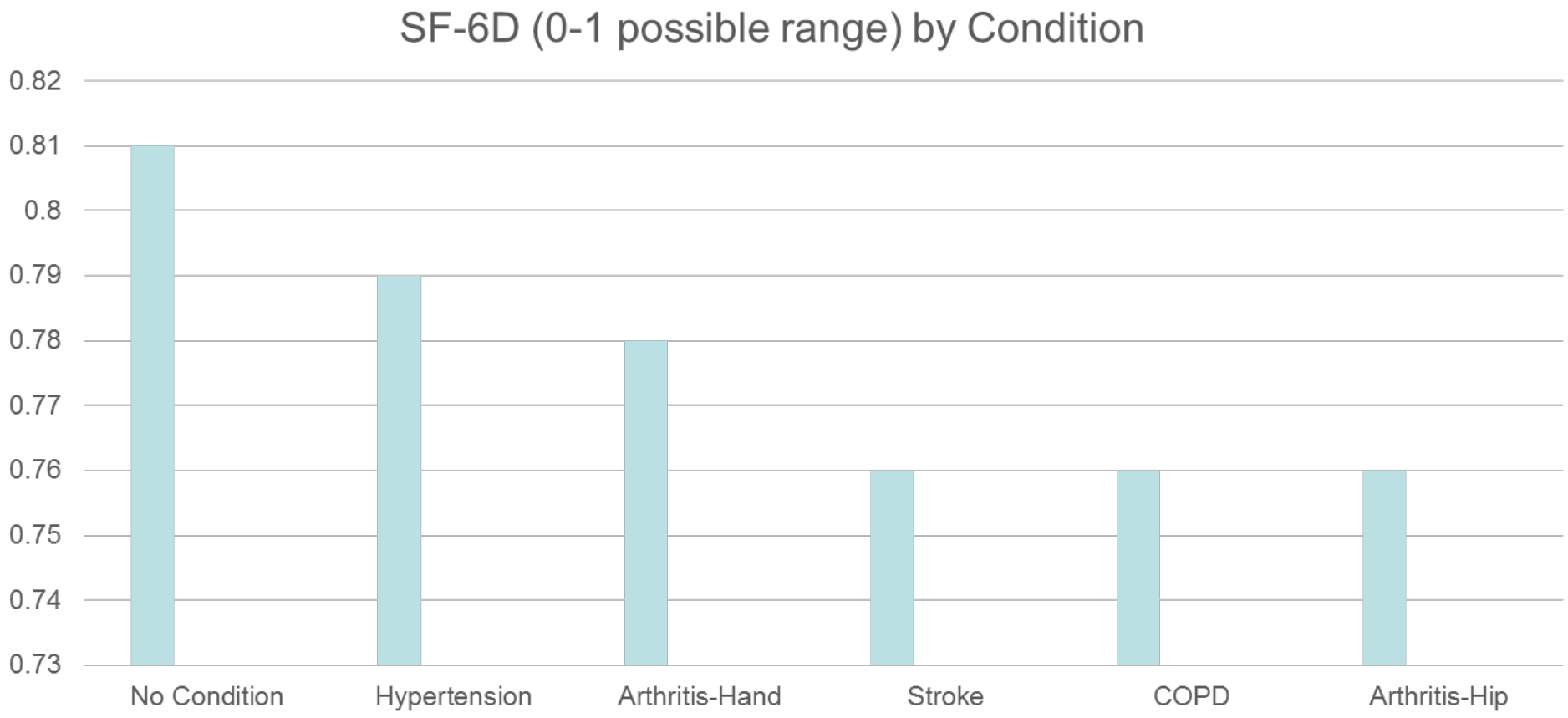
Hays, R.D., Wells, K.B., Sherbourne, C.D., Rogers, W., & Spritzer, K. (1995). Functioning and well-being outcomes of patients with depression compared to chronic medical illnesses. Archives of General Psychiatry, 52, 11-19.

Physical Functioning in Relation to Time Spent Exercising 2-years Before



Stewart, A.L., Hays, R.D., Wells, K.B., Rogers, W.H., Spritzer, K.L., & Greenfield, S. (1994). Long-term functioning and well-being outcomes associated with physical activity and exercise in patients with chronic conditions in the Medical Outcomes Study. Journal of Clinical Epidemiology, 47, 719-730.

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

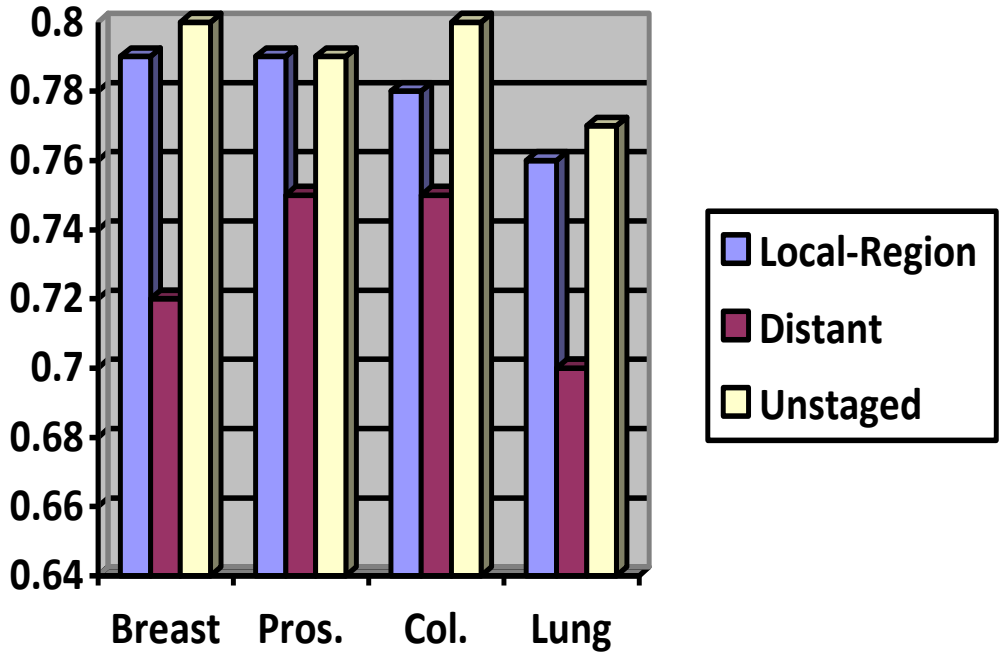
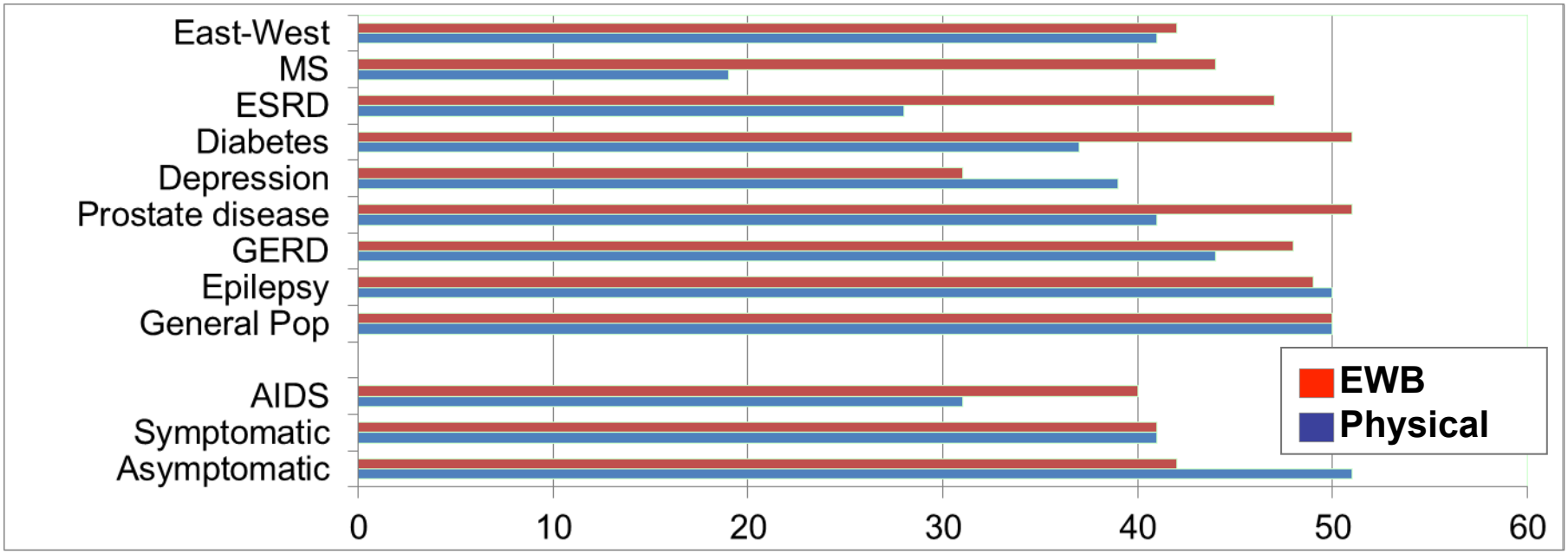


Figure 1. Distant Stage of Disease Associated with Worse SF-6D Scores (Sample sizes for local/regional, distant, and unstaged: Breast (2045,26, 347); Prostate (2652, 61 and 633), Colorectal (1481, 48 and 203), and Lung (466, 47 and 65).

Physical Functioning and Emotional Well-Being at Baseline for 54 Patients at UCLA-Center for East West Medicine



MS = multiple sclerois; 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)

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

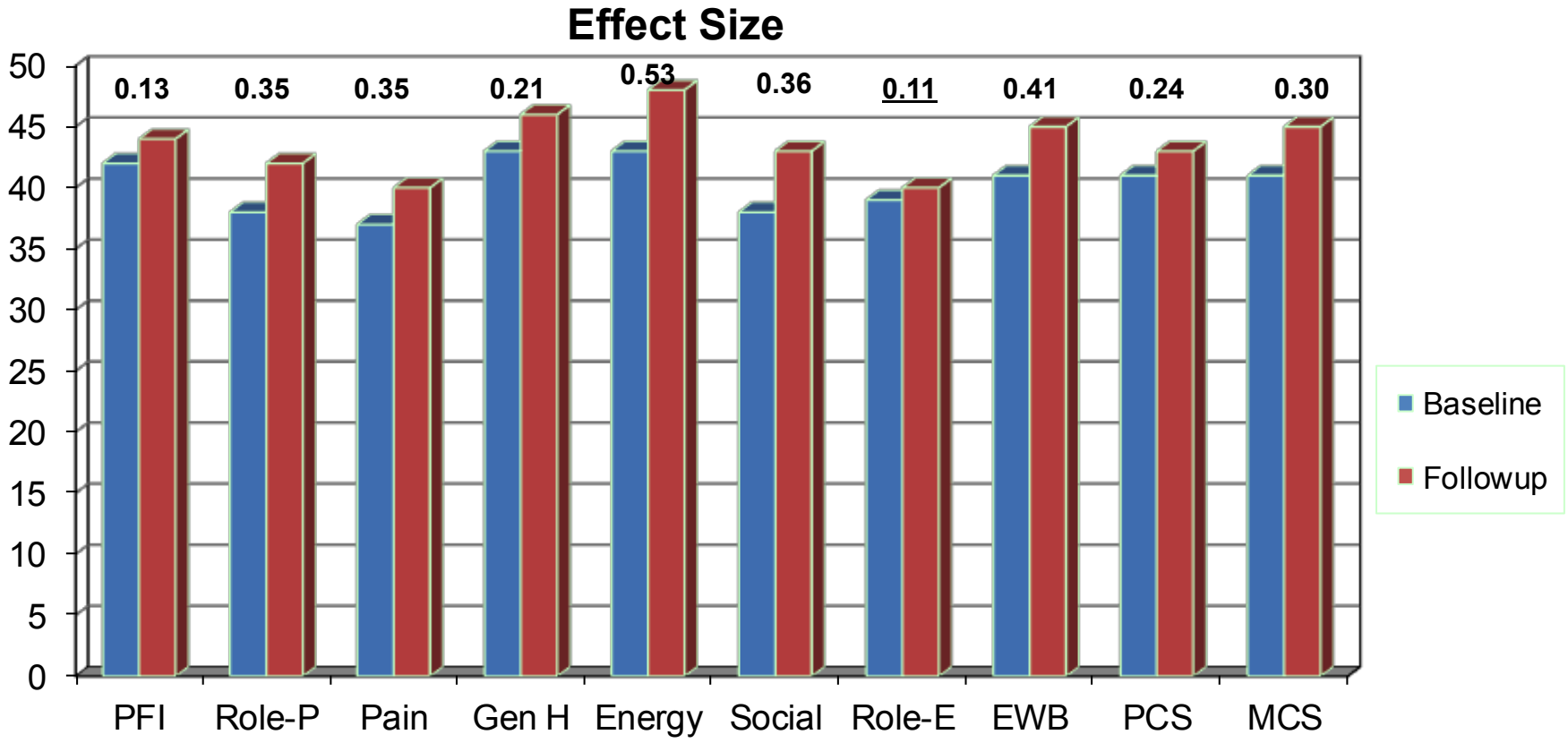
Effect Size

(Follow-up – Baseline)/ $SD_{baseline}$

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



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.

Defining a Responder: Reliable Change Index (RCI)

$$\frac{X_2 - X_1}{(\sqrt{2})(SEM)}$$

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

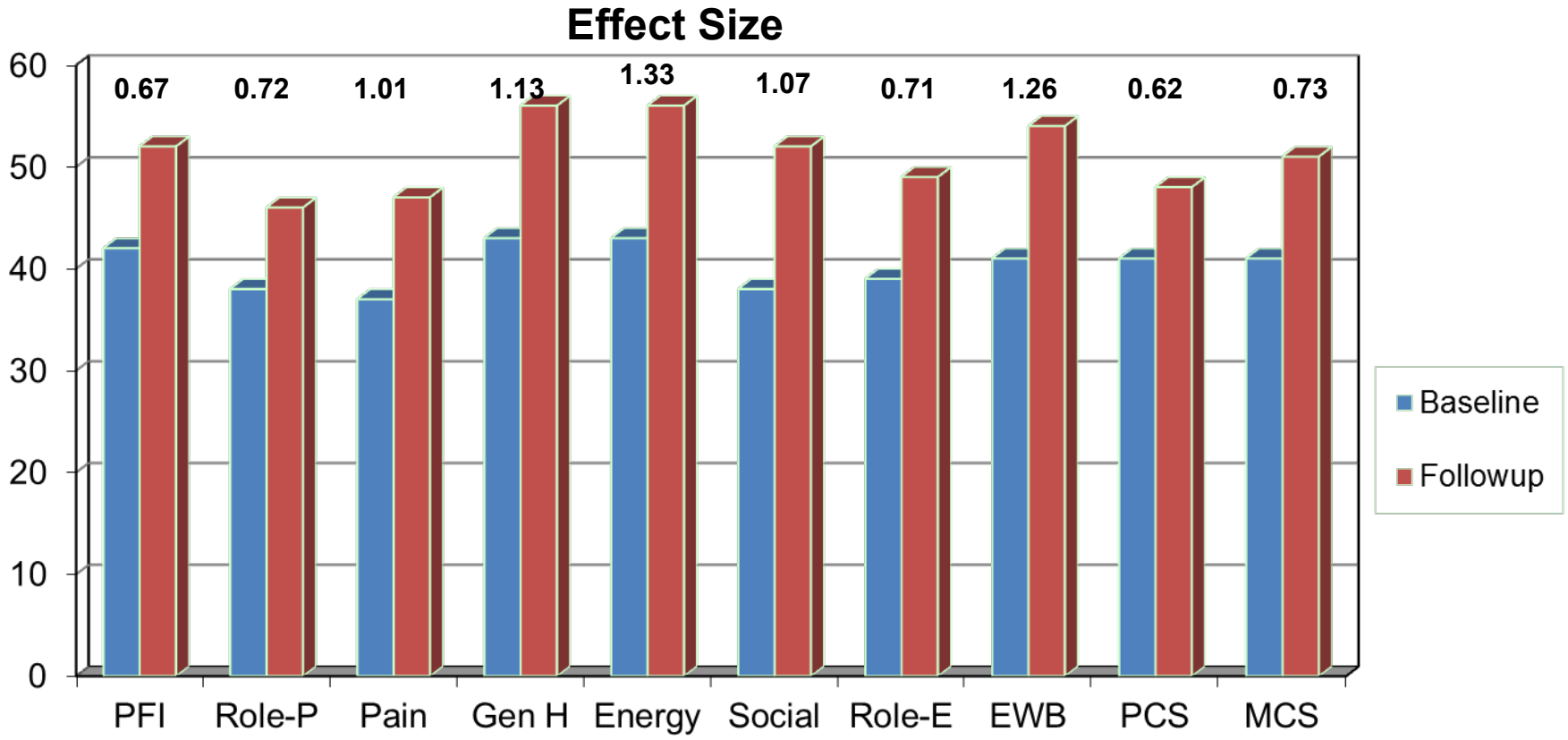
Note: SD_{bl} **= standard deviation at baseline**
r_{xx} **= reliability**

Amount of Change in Observed Score Needed To be Statistically Significant

$$(\sqrt{2})\,(SD_{bl})\sqrt{(1 - r_{xx})}\,(1.96)$$

Note: ***SD_{bl}*** **= standard deviation at baseline and** ***r_{xx}*** **= reliability**

Amount of Change Needed for Significant Individual Change

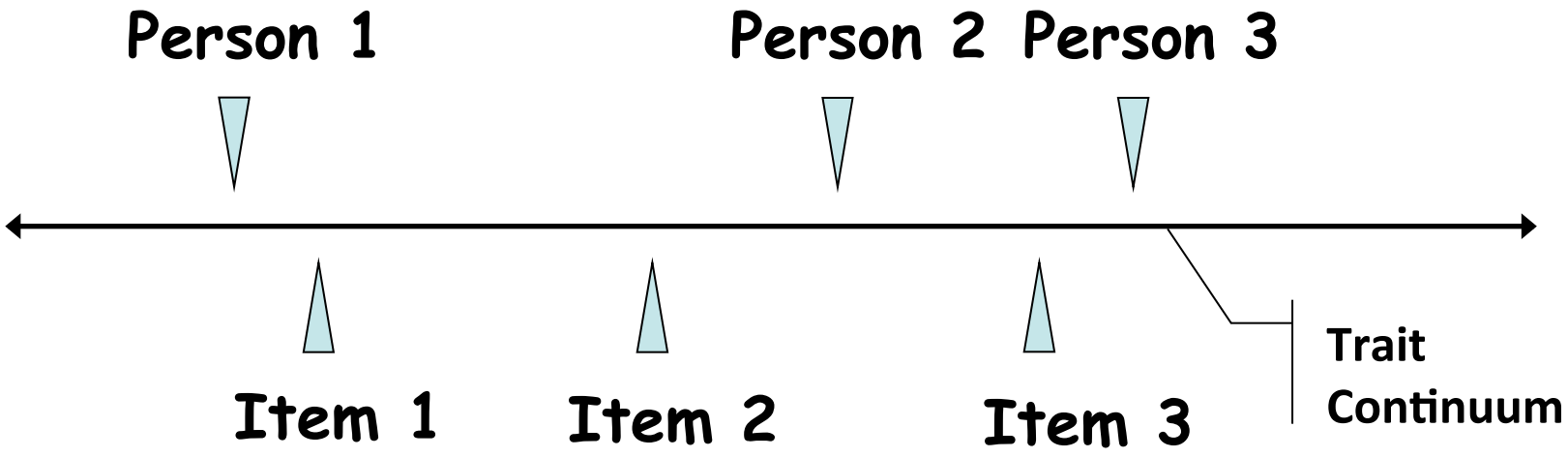


PFI = Physical Functioning; Role-P = Role-Physical; Pain = Bodily Pain; Gen H=General Health; Energy = Energy/Fatigue; Social = Social Functioning; Role-E = Role-Emotional; EWB = Emotional Well-being; PCS = Physical Component Summary; MCS =Mental Component Summary.

7-31% of People in Sample Improve Significantly

	% Improving	% Declining	Difference
PF-10	13%	2%	+ 11%
RP-4	31%	2%	+ 29%
BP-2	22%	7%	+ 15%
GH-5	7%	0%	+ 7%
EN-4	9%	2%	+ 7%
SF-2	17%	4%	+ 13%
RE-3	15%	15%	0%
EWB-5	19%	4%	+ 15%
PCS	24%	7%	+ 17%
MCS	22%	11%	+ 11%

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 $(1 - \text{reliability})^{1/2}$
  - 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)
  - Reliability = $1 - (SE/10)^2$

$$T = 50 + (z * 10)$$

Reliability and SEM

- For z-scores (mean = 0 and SD = 1):
  - Reliability = $1 - SE^2$
  - 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 …

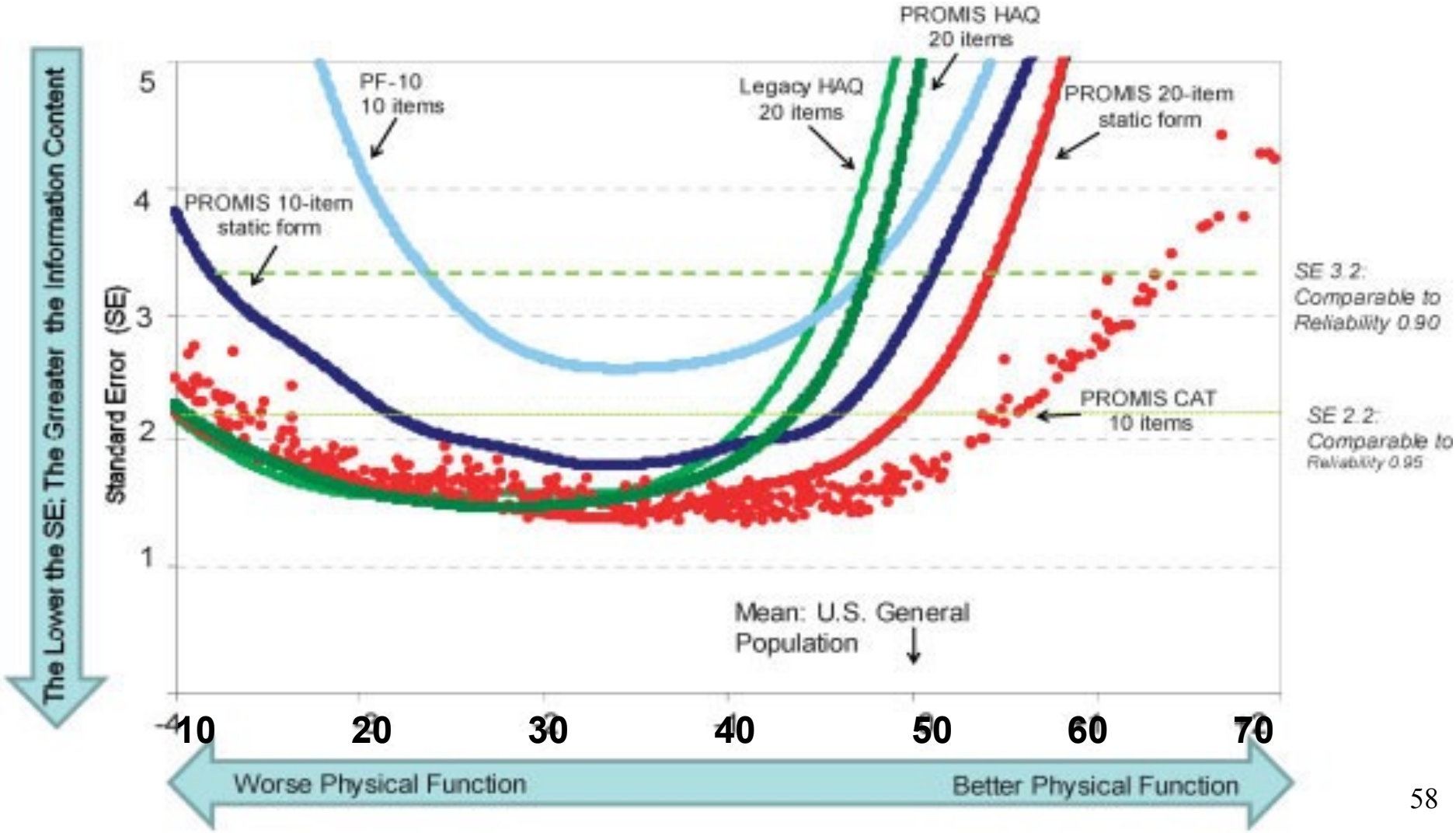
I made myself angry about something just by thinking about it. [6th question]

- Never
- Rarely
- Sometimes
- Often
- Always

Estimated Anger = 50.2

SE = 2.8 (rel = 0.92)

PROMIS Physical Functioning vs. "Legacy" Measures



"Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations"

- Snyder, C.F., Aaronson, N. K., et al. Quality of Life Research, 21, 1305-1314, 2012.

  - 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

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.