

Health-Related Quality of Life - 2008

John E. Ware, Jr., PhD

QualityMetric Incorporated, Lincoln, RI
Tufts University School of Medicine, Boston, MA

Introduction to the Principles and Practice of Clinical Research
National Institutes of Health
Clinical Center
Building 10, Lipsett Amphitheater, Bethesda, MD
January 28, 2008



Quality of Life

Community
Education
Family Life
Friendships
Health
Housing

Marriage
Nation
Neighborhood
Self
Standard of
Living
Work

Source: Campbell, 1981

What is Your Health?

- Bodily structure & function
- What you are able to do – functioning
- How you feel – distress & well-being
- What you say it is – personal evaluation

Sources: Understanding Health Outcomes Educational Series

Background

- Patient-reported outcomes (PROs) are very useful
- Standardization of concepts & metrics is enabling comparisons & interpretation across applications
- Widespread use proves that more practical short forms will be adopted
- Traditional “static” short forms rarely meet clinical standards of precision
- Advances, including item response theory (IRT), computerized adaptive testing (CAT) and the Internet may provide solutions

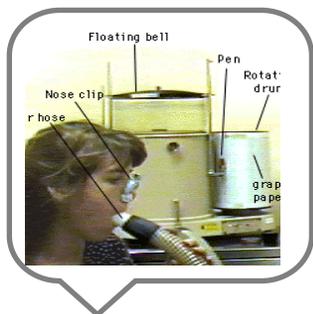
Continuum of Disease-specific and General Health Measures



Adapted from: Wilson and Cleary, *JAMA*, 1995
Ware, *Annual Rev. Pub. Health*, 1995

Continuum of Disease-specific and General Health Measures

Spirometry



Shortness of Breath

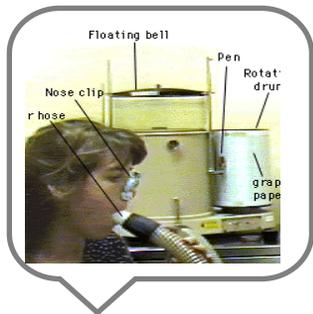
Over the last 4 weeks I have had shortness of breath
Almost every day
Several days a week
A few days a month
Not at all



Adapted from: Wilson and Cleary, *JAMA*, 1995
Ware, *Annual Rev. Pub. Health*, 1995

Continuum of Disease-specific and General Health Measures

Spirometry



Shortness of Breath

Over the last 4 weeks I have had shortness of breath
Almost every day
Several days a week
A few days a month
Not at all

Respiratory-specific

How much did your lung/respiratory problems limit your usual activities or enjoyment of everyday life?
Not at all
A little
Moderately
Extremely

Clinical Markers

Specific Symptoms

Impact of Disease-specific Problems

Generic Functioning, Well-being and Evaluation

(1)

(2)

(3)

(4)

Adapted from: Wilson and Cleary, *JAMA*, 1995
Ware, *Annual Rev. Pub. Health*, 1995

Which Is Most Valid For the Purposes of Measuring Health Outcome

Shortness of Breath

Over the last 4 weeks I have had shortness of breath
Almost every day
Several days a week
A few days a month
Not at all

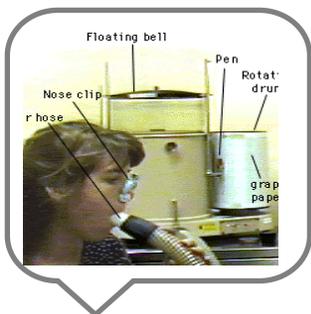
Respiratory-specific

How much did your lung/respiratory problems limit your usual activities or enjoyment of everyday life?
Not at all
A little
Moderately
Extremely

Adapted from: Wilson and Cleary, *JAMA*, 1995
Ware, *Annual Rev. Pub. Health*, 1995

Continuum of Disease-specific and General Health Measures

Spirometry



Shortness of Breath

Over the last 4 weeks I have had shortness of breath

Almost every day
 Several days a week
 A few days a month
 Not at all

Respiratory-specific

How much did your lung/respiratory problems limit your usual activities or enjoyment of everyday life?

Not at all
 A little
 Moderately
 Extremely

Generic

SF-36® Health Survey Scales, Summary Measures



(1)



(2)



(3)



(4)

Adapted from: Wilson and Cleary, *JAMA*, 1995
 Ware, *Annual Rev. Pub. Health*, 1995

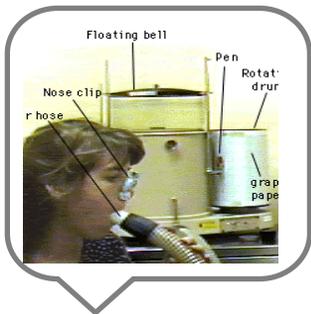
Continuum of General Health

Most frequently used generic survey according to Garratt et al. Quality of Life Measurement Bibliographic Study. *BMJ*, 2002;324:1417-22

Spirometry

Shortness of Breath

Generic



Over the last 4 weeks I have had shortness of breath
 Almost every day
 Several days a week
 A few days a month
 Not at all

How much did your lung/respiratory problems limit your usual activities or enjoyment of everyday life?
 Not at all
 A little
 Moderately
 Extremely

SF-36® Health Survey Scales, Summary Measures



(1)

(2)

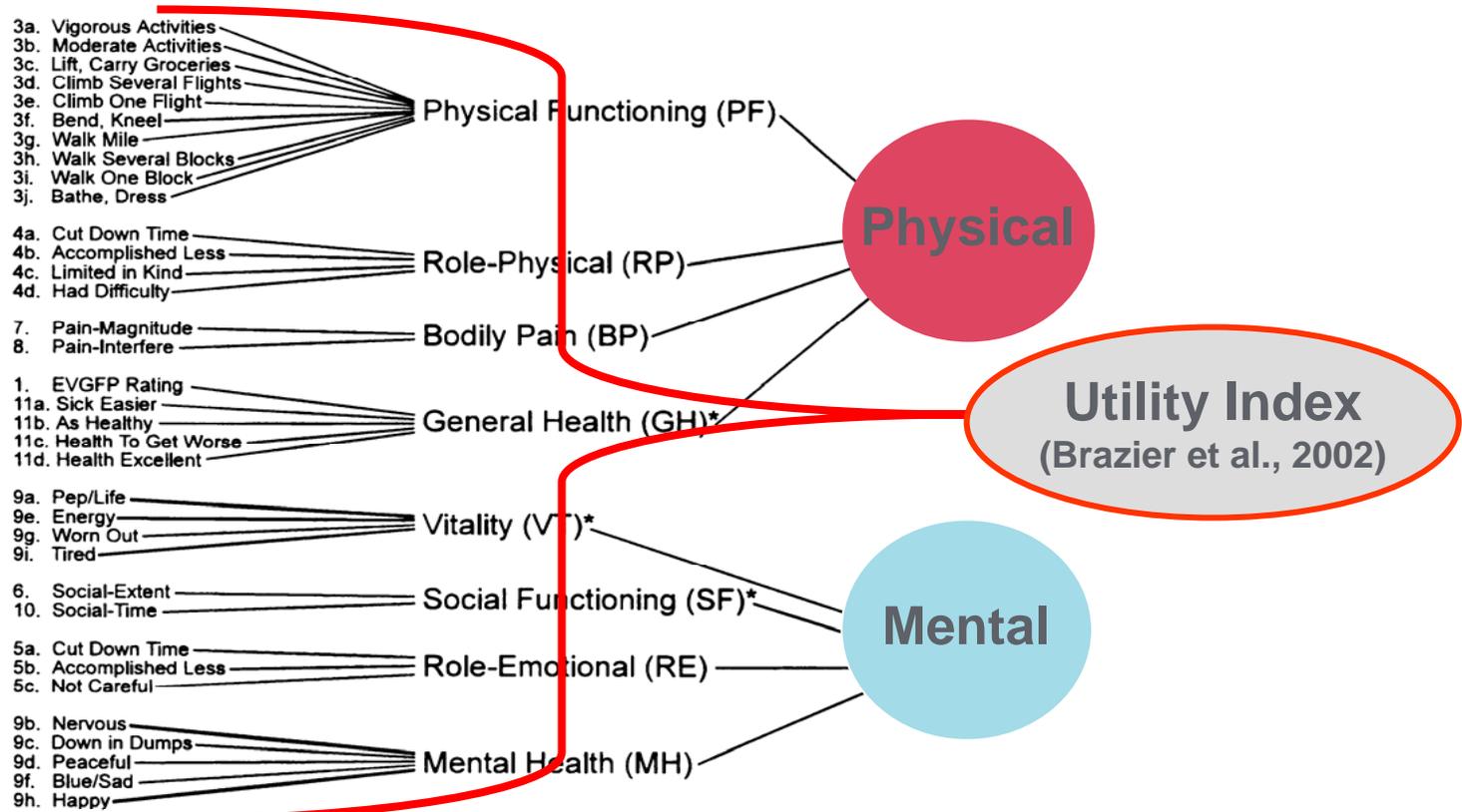
(3)

(4)

Health-related QOL

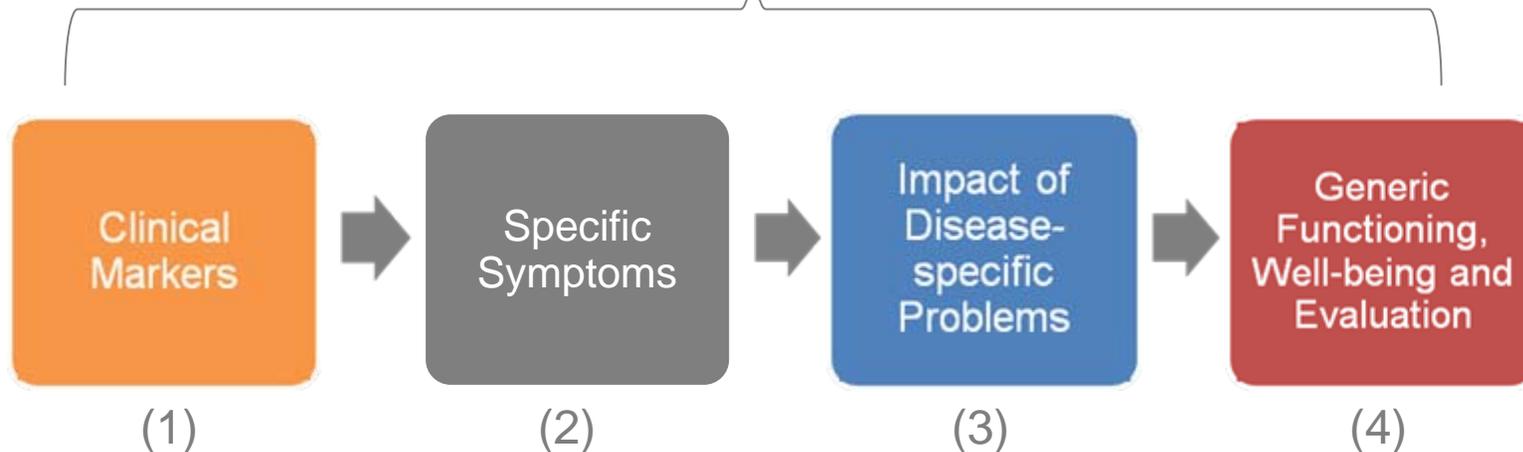
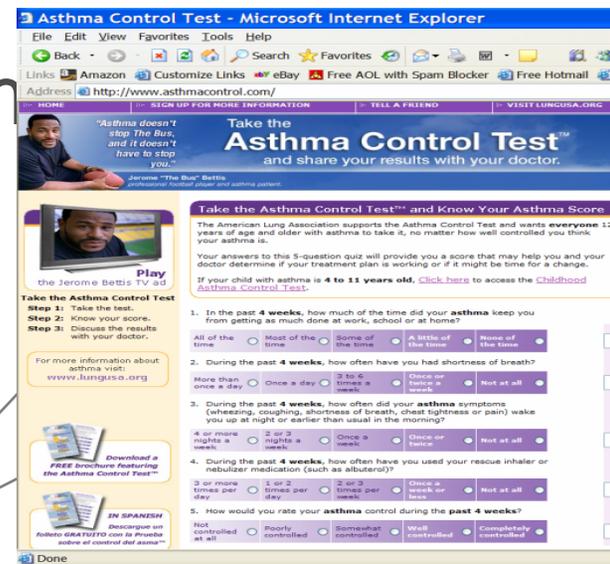
Adapted from: Wilson and Cleary, *JAMA*, 1995
 Ware, *Annual Rev. Pub. Health*, 1995

SF-36[®] Health Survey Measurement Model



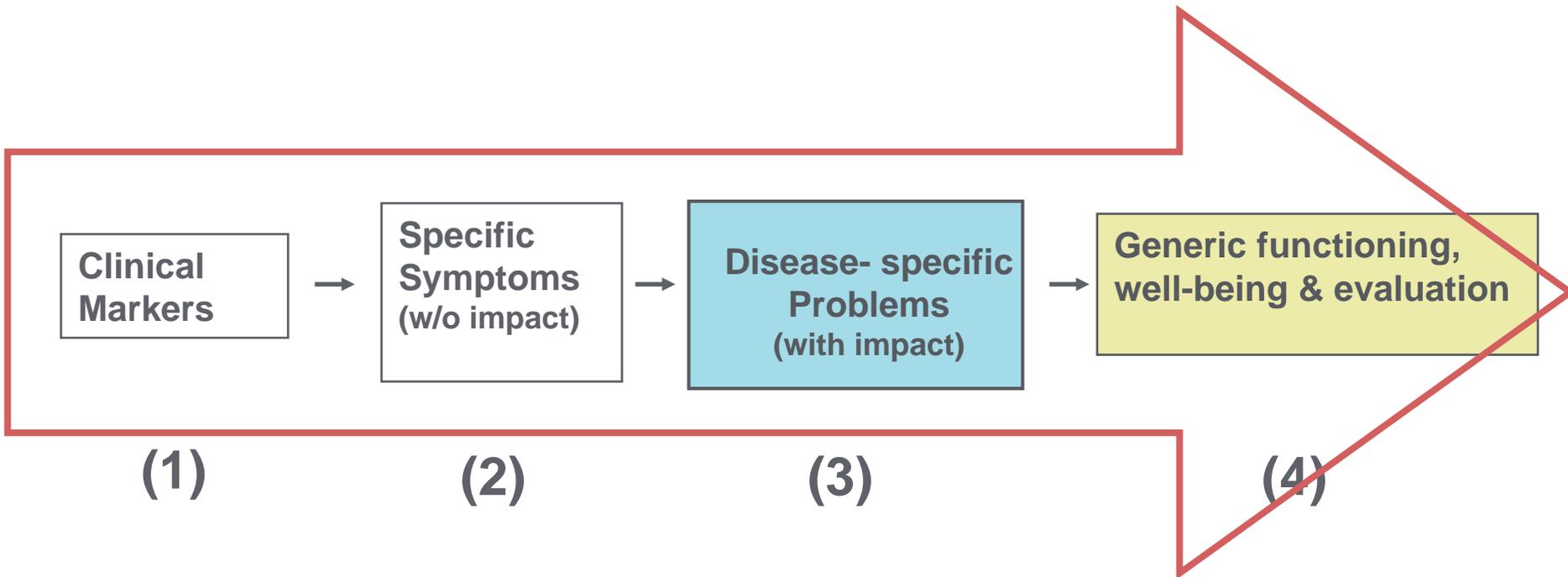
* Significant correlation with other summary measure.

Comprehensive Outcome Asthma Control Test

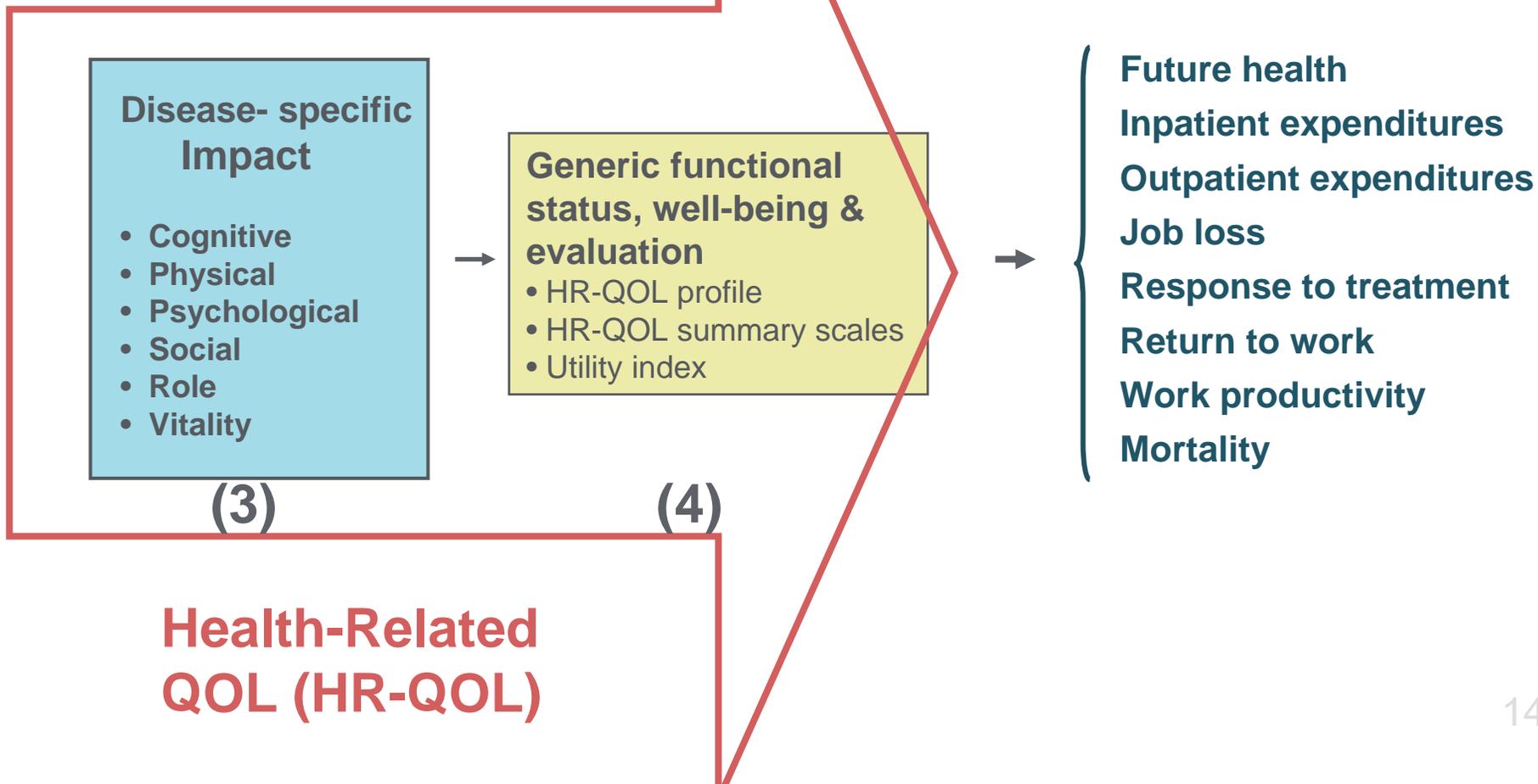


References: Nathan RA, Sorkness CA, Kosinski M et al., "Development of the Asthma Control Test: A survey for assessing asthma control." *Journal of Allergy and Clinical Immunology*. 2004;113:59-65.
 Schatz M, Sorkness CA, Li JT, Marcus P, et al, "Asthma Control Test: Reliability, validity, and responsiveness in patients not previously followed by asthma specialists," *Journal of Allergy and Clinical Immunology*, 2006;117:549-556.

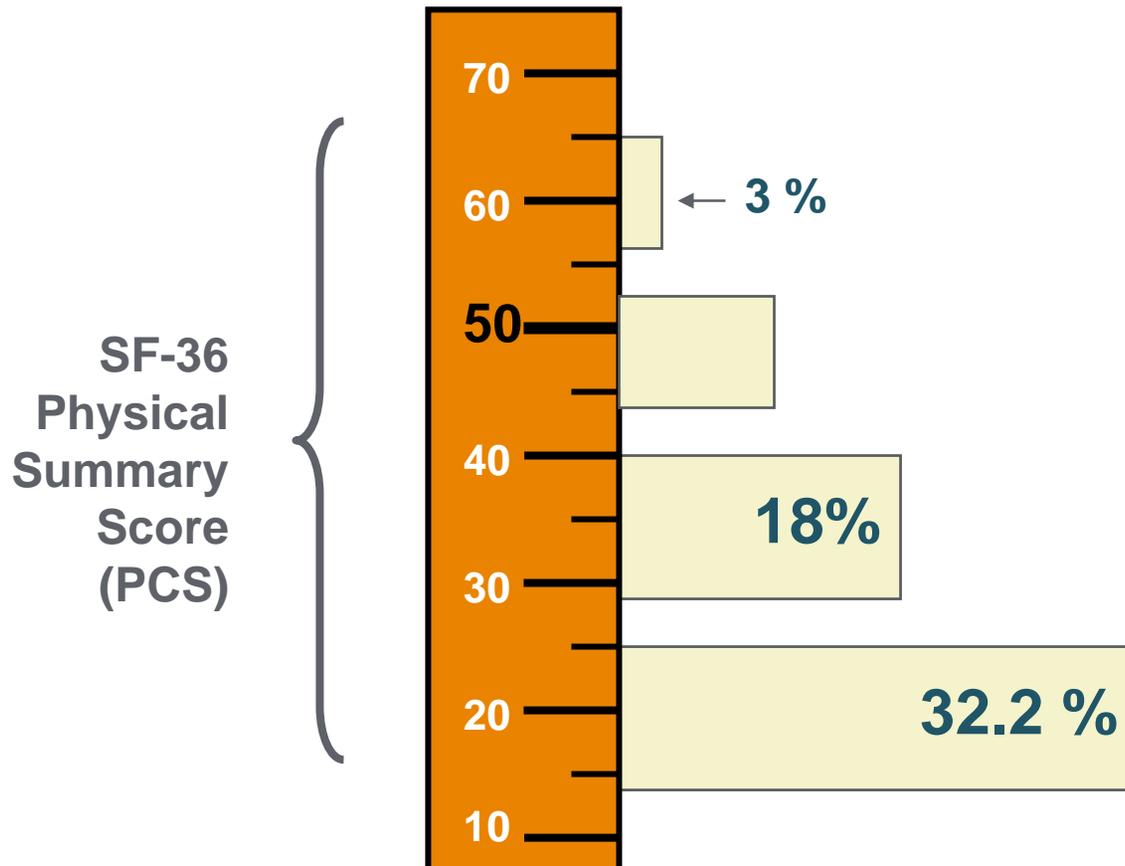
There is More to the Continuum



Prediction and Risk Management: HRQOL in one of the Best Predictors

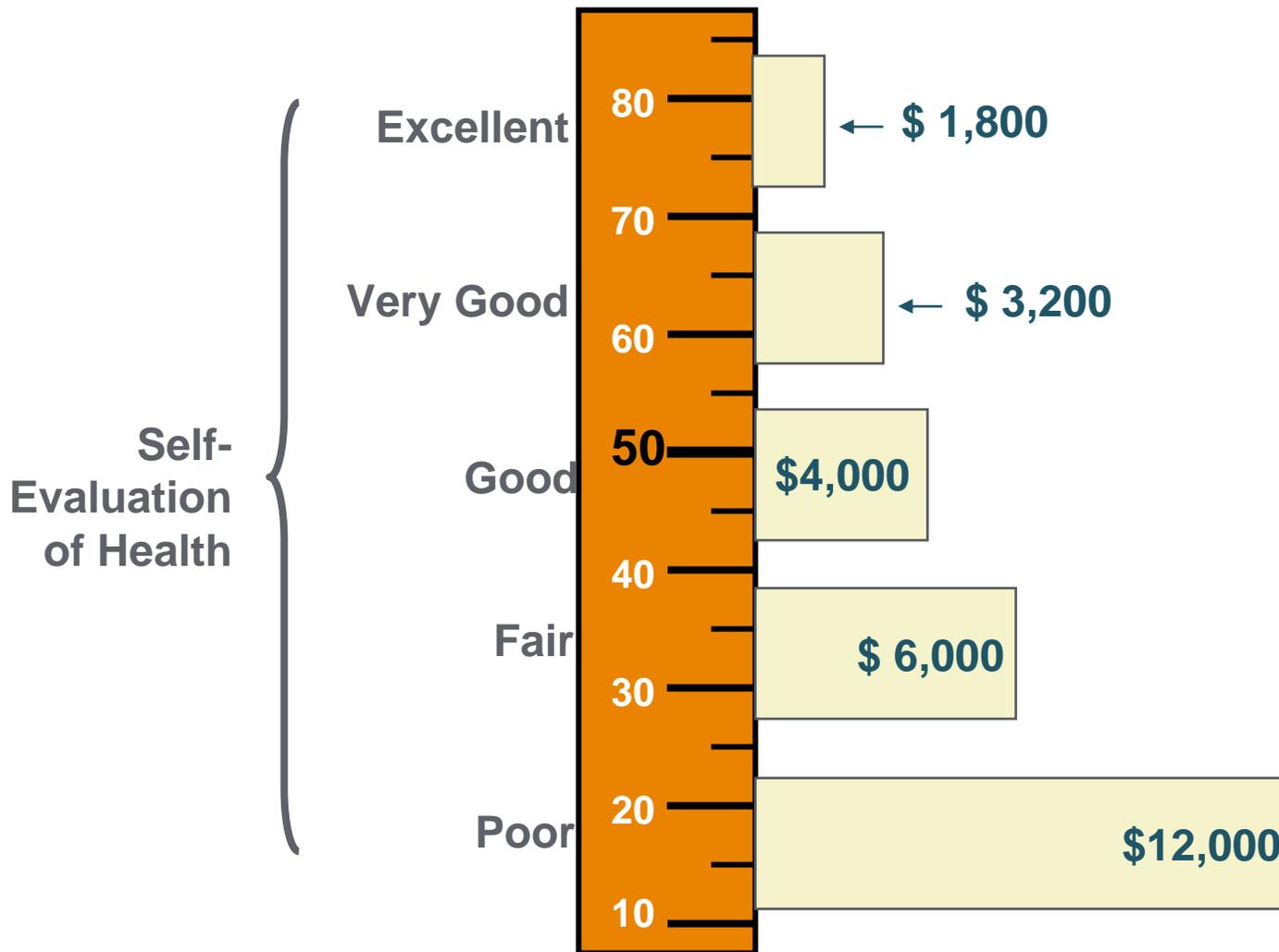


Physical Health Status Predicts: Job Loss Due to Health Problems

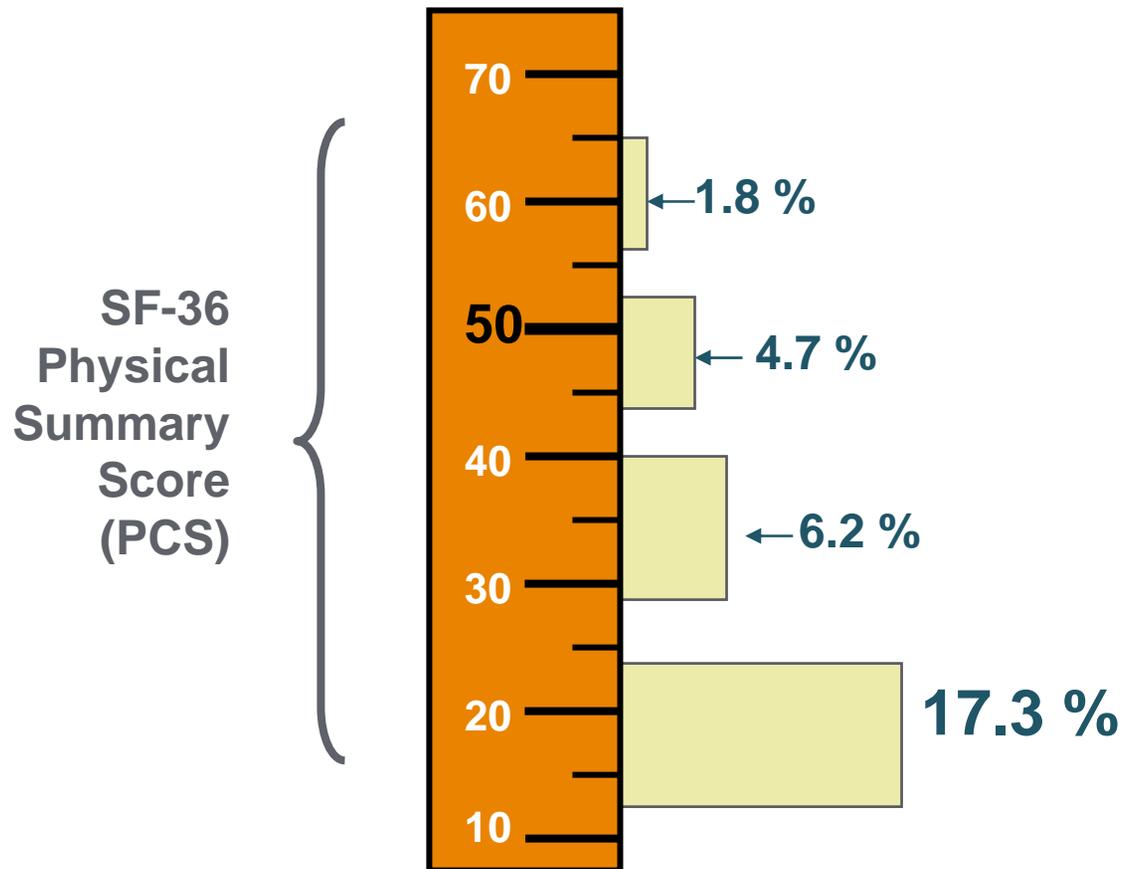


Source: Ware, Kosinski and Keller, 1994

Self-Assessed Health Status Predicts: Annual Costs of Health Care



Physical Health Status Predicts: 5-Year Mortality (All Causes)





What Do We Need for Clinical Research & Practice?

Medical Outcomes Trust (MOT) Review Criteria

- I. Conceptual & Measurement Model
- II. Reliability
- III. Validity
- IV. Responsiveness
- V. Interpretability
- VI. Burden
- VII. Alternative Forms
- VIII. Cultural & Language Adaptations

Source: MOT Scientific Advisory Committee, Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research*. 2002 11 (3): 193-205.

What Do We Need for Clinical Research and Practice?

- **Outcomes that matter to patients**
- **Practical measures**
- **Coverage of a wide range**
- **Greater precision**
- **Comparability of scores**
- **Ease of interpretation**

Summary of Information About Widely-Used General Health Scales

NIH Roadmap PROMIS Initiative:

- Fatigue
- Negative affect
- Pain
- Physical Function
- Social/role activity

www.nihpromis.org

Concepts and Characteristics	Psychometric									
	SIP	HIE	NHP	QLI	COOP	DUKE	HUI	SF-6D	QWB	MOS SF-36
CONCEPTS										
Physical functioning	•									
Social functioning	•	•	•	•	•	•	•	•	•	•
Role functioning	•	•	•	•	•	•	•	•	•	•
Psychological distress	•	•	•	•	•	•	•	•	•	•
Health perceptions (general)		•	•	•	•	•	•	•	•	•
Pain (bodily)		•	•	•	•	•	•	•	•	•
Energy/fatigue		•	•	•	•	•	•	•	•	•
Psychological well-being		•	•	•	•	•	•	•	•	•
Sleep	•		•			•	•			
Cognitive functioning	•					•	•			•
Quality of life		•				•	•			
Reported health transition						•	•	•		

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 and 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)

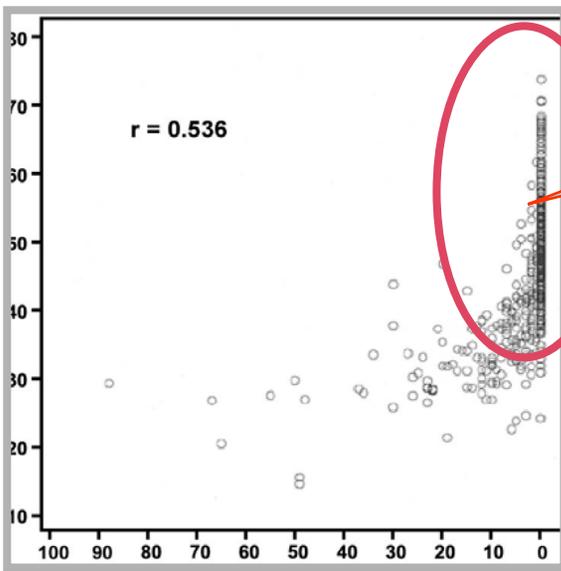
HUI = Health Utility Index (1996)

SF-6D = SF-36 Utility Index (Brazier, 2002)



What Do We Need for Clinical Research and Practice?

- Outcomes that matter to patients
- **Practical measures**
- Coverage of a wide range
- Greater precision
- Comparability of scores
- Ease of interpretation



“Ceiling Effect”

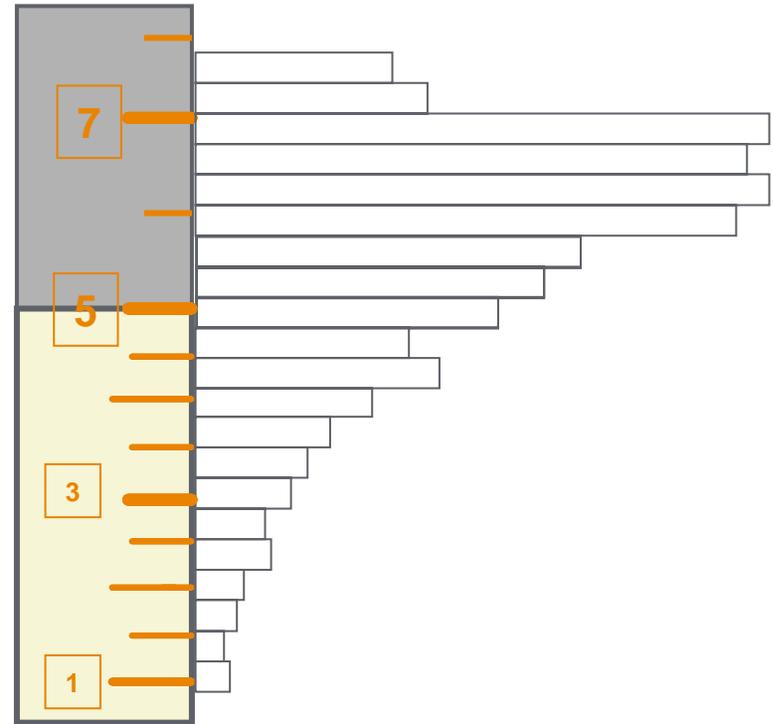
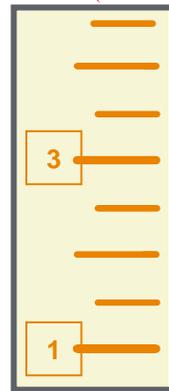
What Do We Need for Clinical Research and Practice?

- Outcomes that matter to patients
- Practical measures
- **Coverage of a wide range**
- Greater precision
- Comparability of scores
- Ease of interpretation

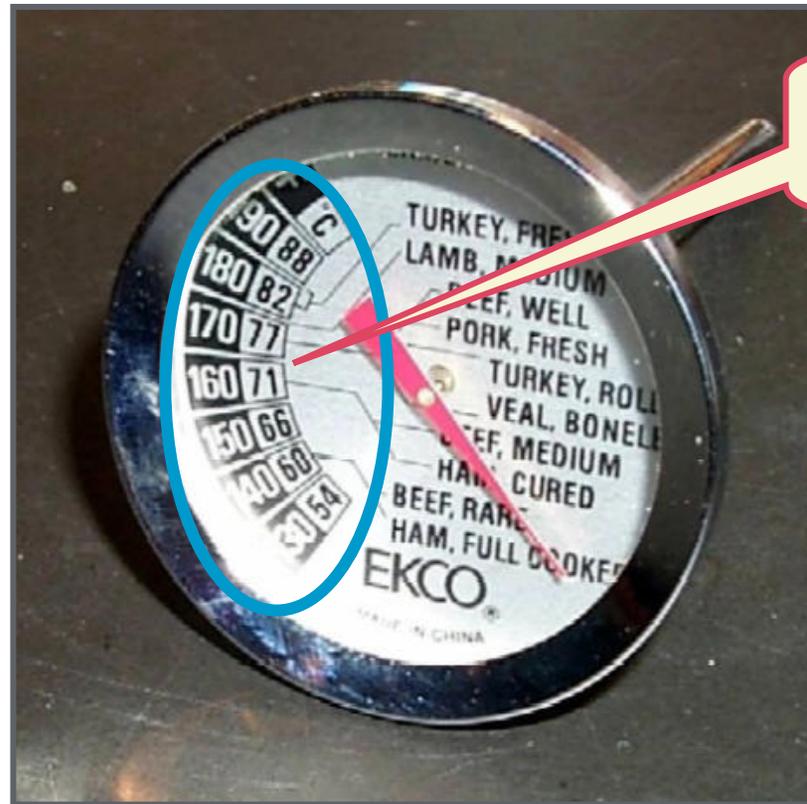
Short-Form Surveys

Do Not Provide a Solution

**Measuring
Too Low -
Ceiling
Effect**



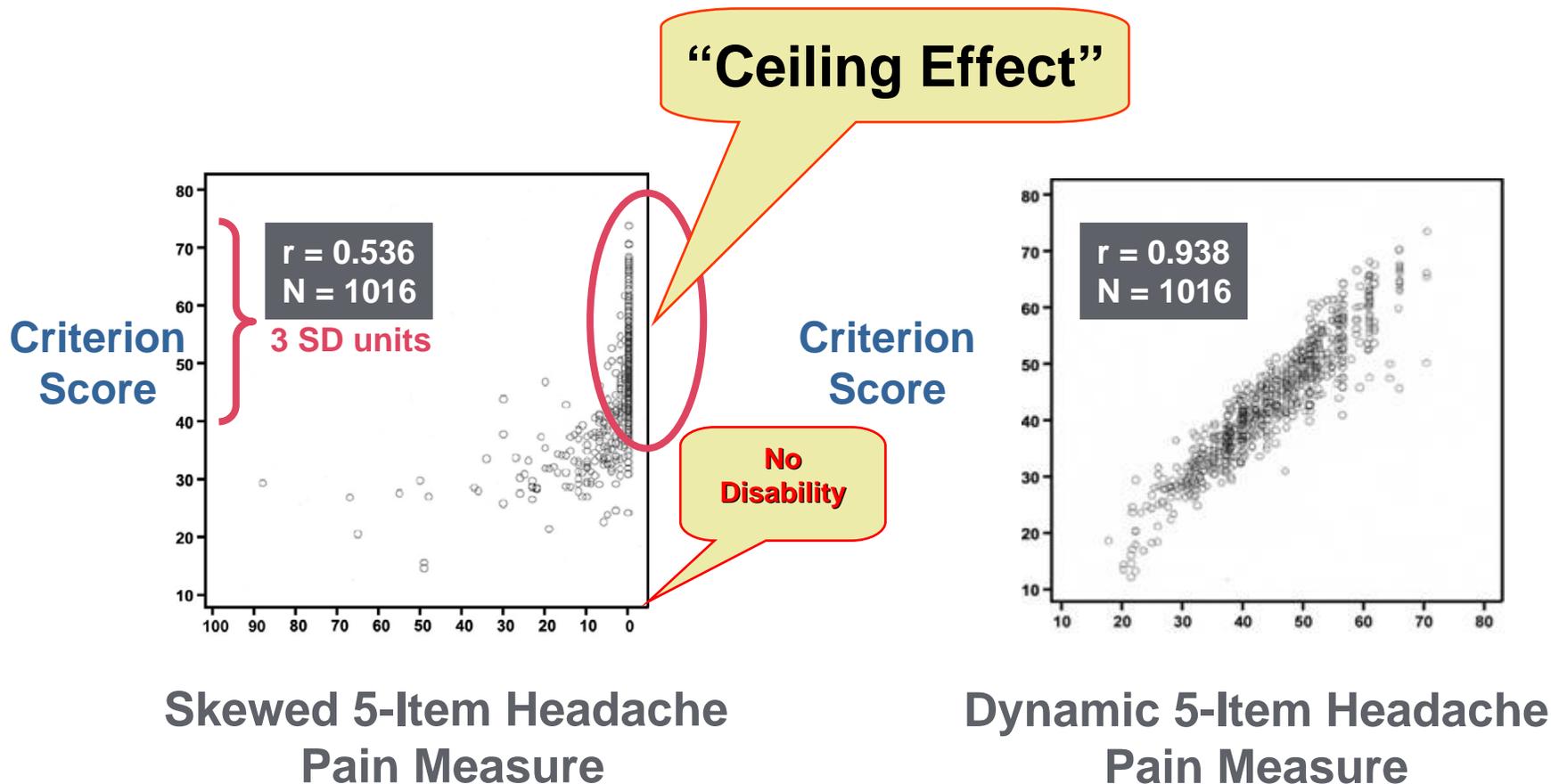
Some Thermometers Focus on a Very Narrow Range

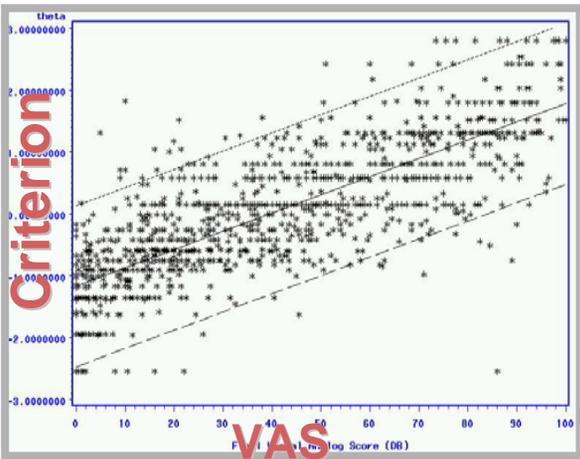


130–190 °F
54–88 °C

Cooking Thermometer

Another Solution is Computerized Dynamic Health Assessment





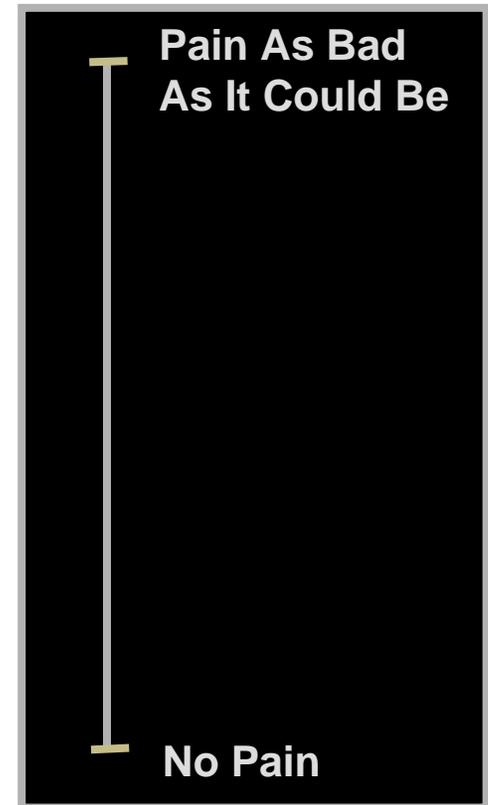
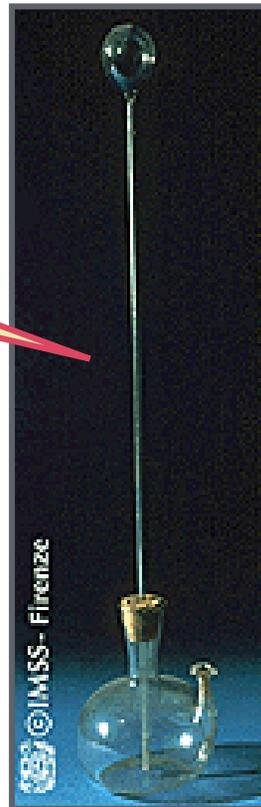
What Do We Need for Clinical Research and Practice?

- Outcomes that matter to patients
- Practical measures
- Coverage of a wide range
- **Greater precision**
- Comparability of scores
- Ease of interpretation

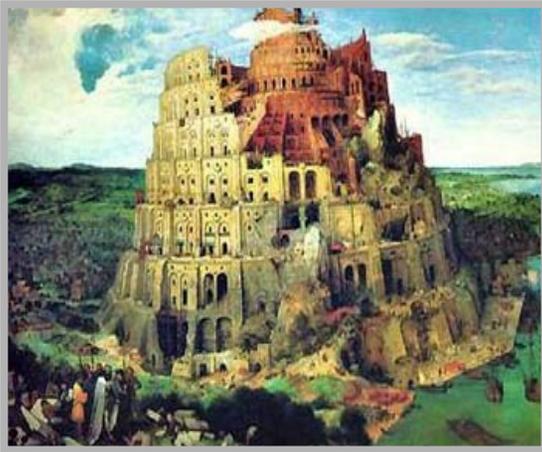
Original Thermoscope

Results were not interpretable

- No marks on the “ruler”
- Poor reproducibility
- No interpretation guidelines



Visual Analogue Scale (VAS)

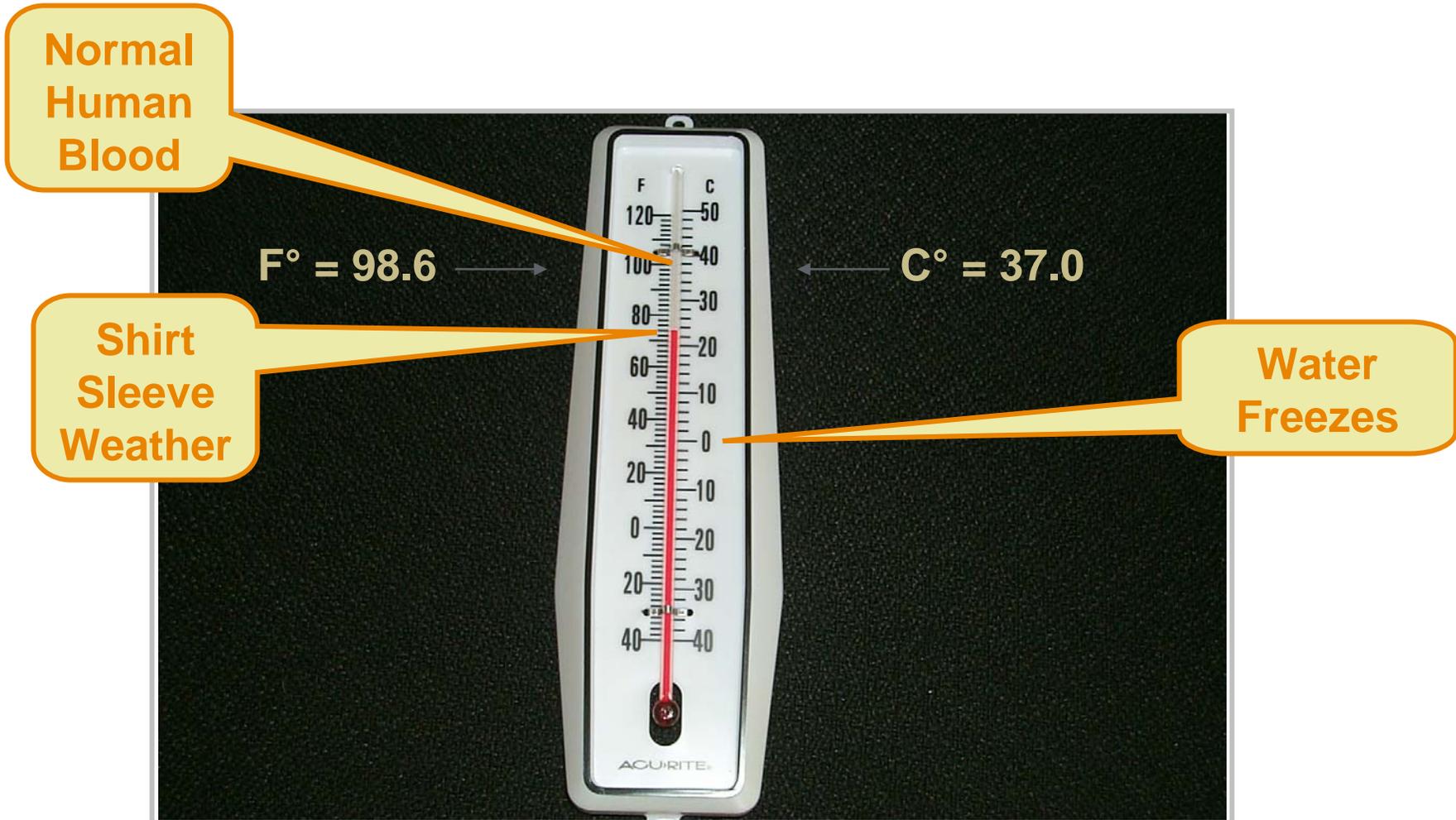


What Do We Need for Clinical Research and Practice?

- Outcomes that matter to patients
- Practical measures
- Coverage of a wide range
- Greater precision
- **Comparability of scores**
- Ease of interpretation

Example: Cross-Calibrating Celsius and Fahrenheit

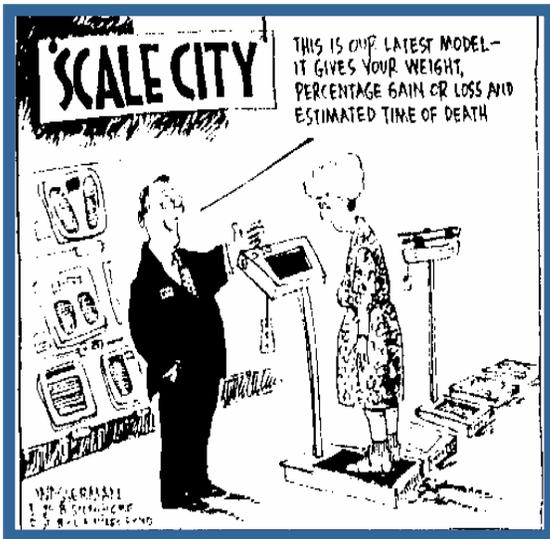
Temperature



We Need the Health Equivalent of a Two-Sided Tape Measure



and Public-Private Partnerships That Meet the Needs of Research and Business



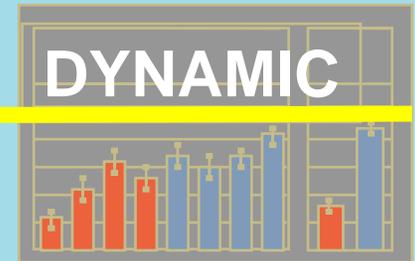
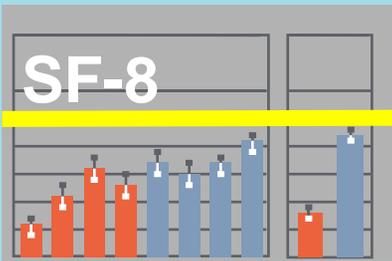
What Do We Need for Clinical Research and Practice?

- Outcomes that matter to patients
- Practical measures
- Coverage of a wide range
- Greater precision
- Comparability of scores
- **Ease of interpretation**

Standardization



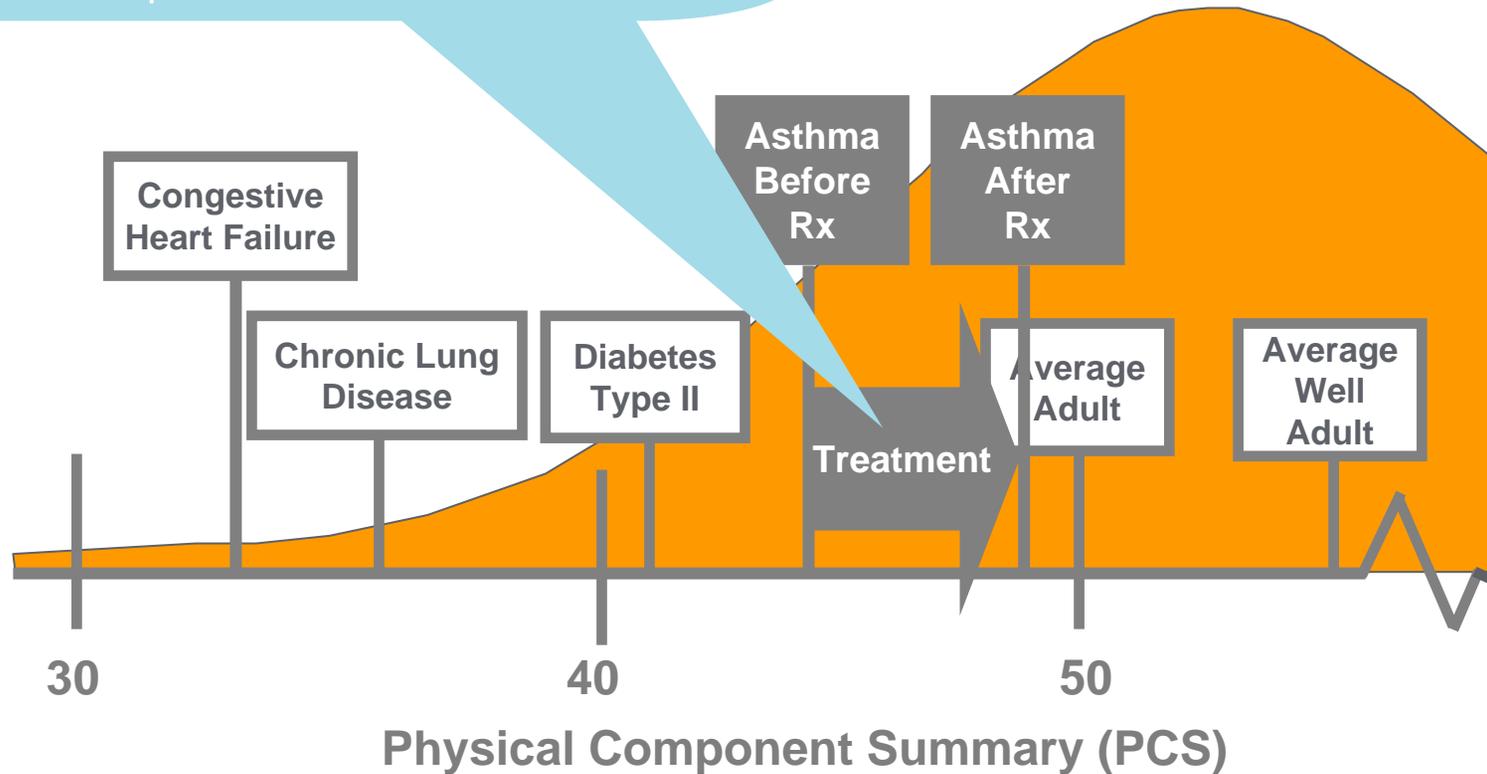
Scoring Software

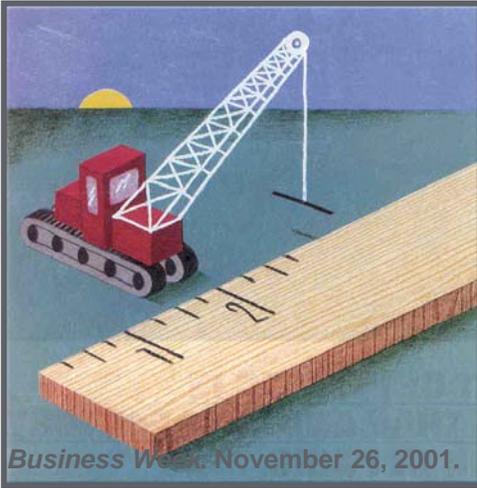


“Improvements in Short Form Measures of Health Status,”
J. Clinical Epidemiology, 2008

What Does It All Mean?

- 50% reduction in physical disability
- 33% reduction in hospitalization
- Substantial increase in work productivity
- Reduced expenditures

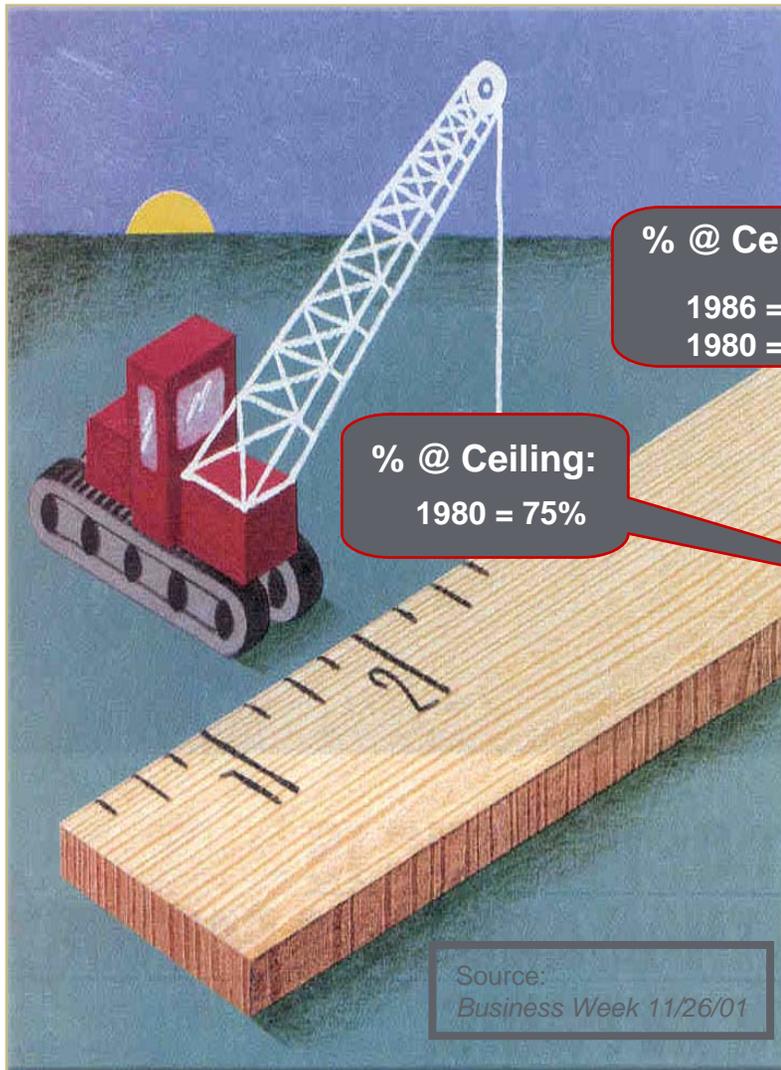




Solutions

- Improved psychometrics
- Computerized adaptive testing (CAT) software
- The Internet (and other connectivity)

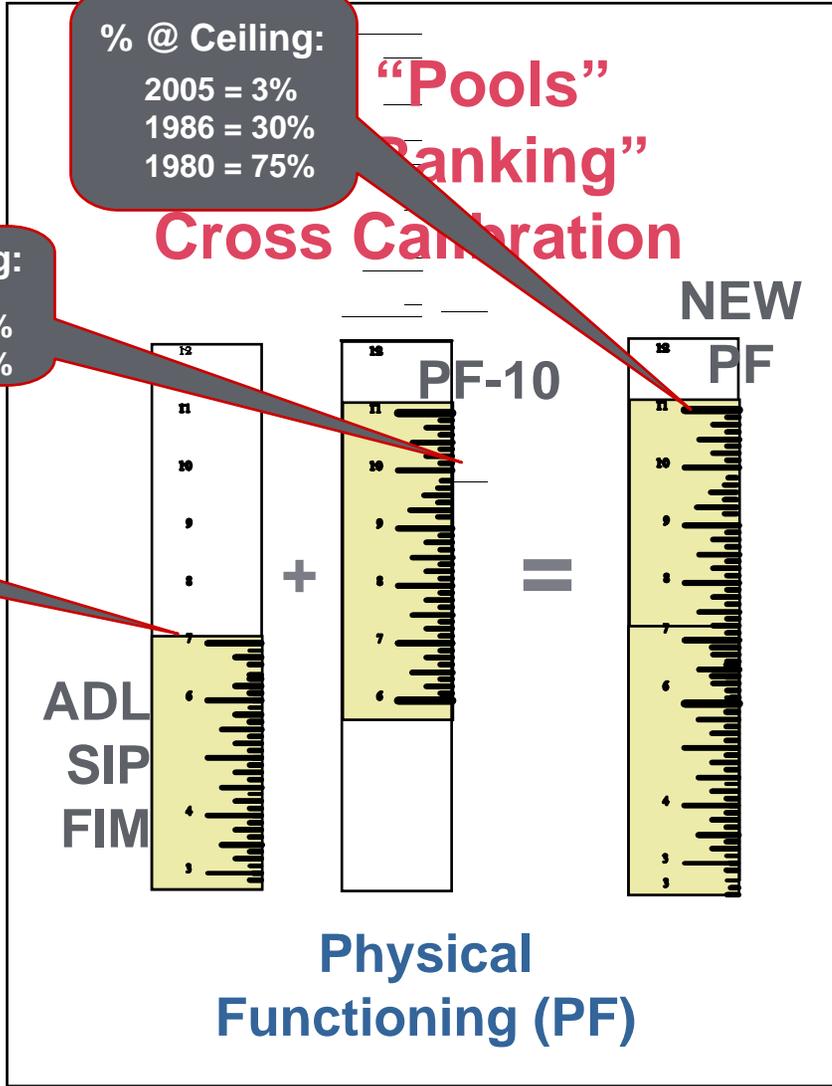
First, Construct Better Metrics



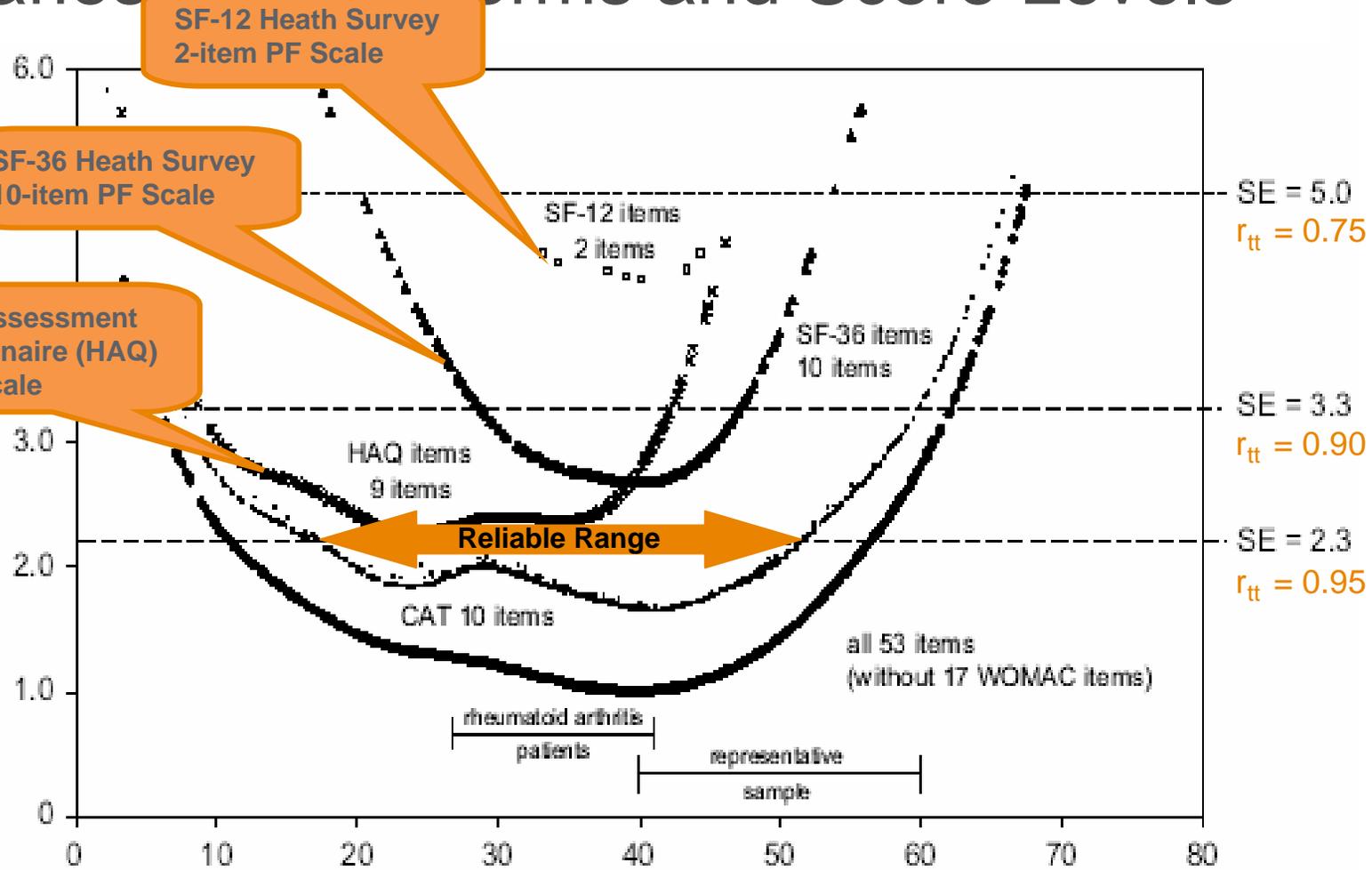
% @ Ceiling:
1980 = 75%

% @ Ceiling:
1986 = 30%
1980 = 75%

% @ Ceiling:
2005 = 3%
1986 = 30%
1980 = 75%



Measurement Precision (Standard Error) Varies Across Forms and Score Levels



normed theta values

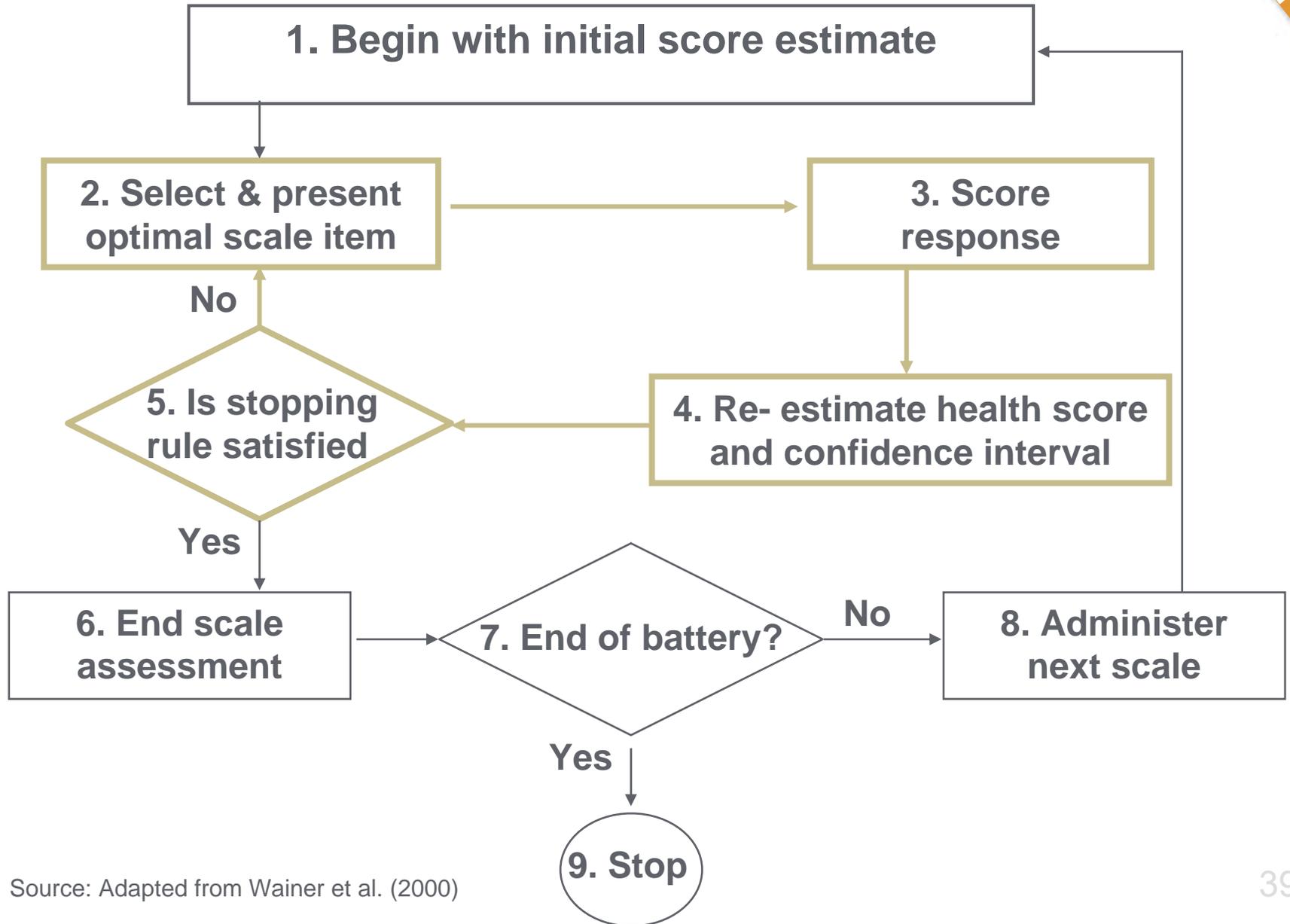
Physical Function Scores (Mean = 50)

2nd Solution, Assess Health Dynamically

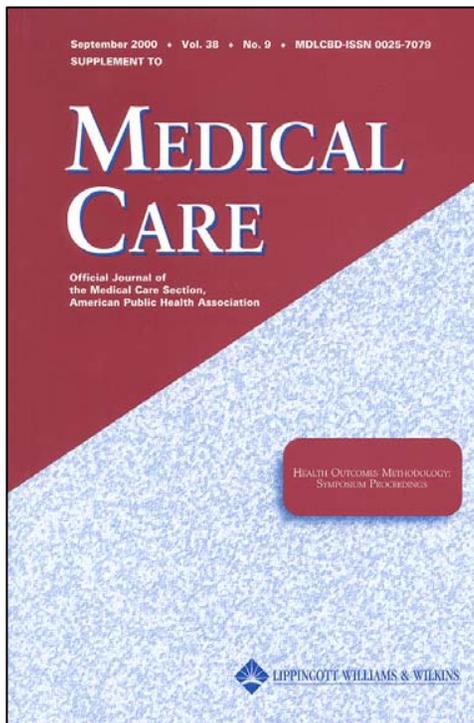
Patient scores here

CAT

Logic of Computerized Adaptive Testing (CAT)



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.

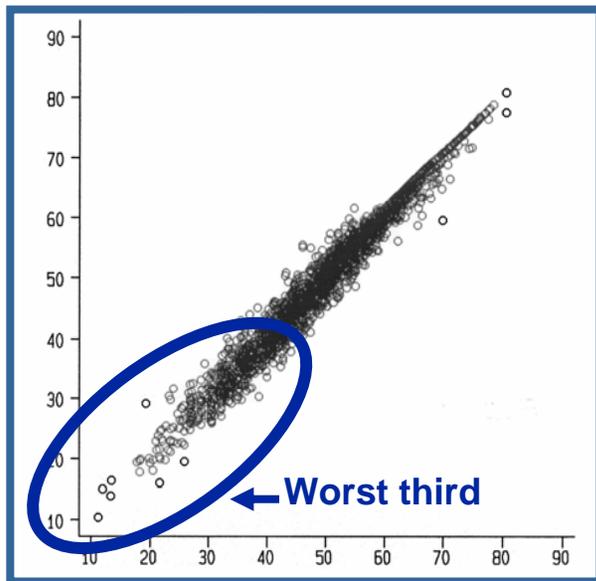
Practical Implications of Item Response Theory and Computerized Adaptive Testing

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



Do CAT Health Assessments Reduce Respondent Burden?

Results (N = 2,753):

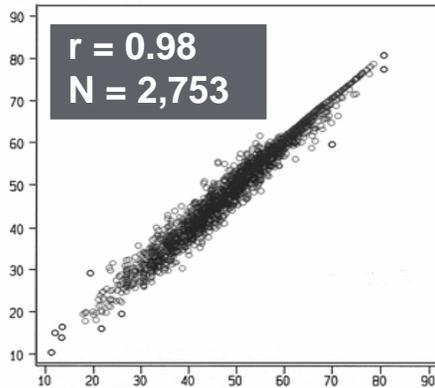
Among patients in poorest mental health (worst third):

92% met clinical standard of precision with five or fewer questions

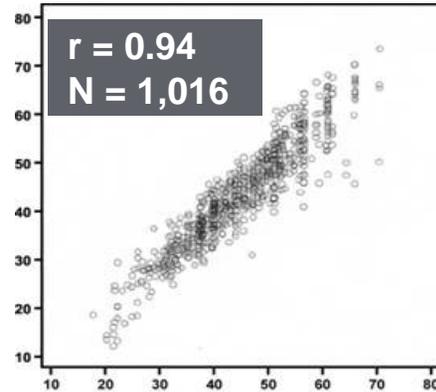
Acknowledgements

Functional Health CAT in Chronic Kidney Disease	NIDDK	Ware, Raczek
Mental Health CAT for Community-Based Use	NIMH	Rose, Ware, Sherbourne
CAT Assessment of Headache Impact	NINDS	Turner-Bowker, Bjorner Ware
CAT Assessment of Asthma Impact	NHLBI/NIAID	Turner-Bowker, Bjorner, Saris-Baglana, Ware
CAT Assessment of Disease Impact	NIA/AHRQ	Bjorner, Ware, Turner- Bowker, Becker
CAT Pediatric Health Assessment	NICHD	Ware, Raczek, Bjorner, Saris-Baglana
CAT Post-Acute Rehabilitation Care	NIDDR/AHRQ	Ware, Bjorner, Gandek

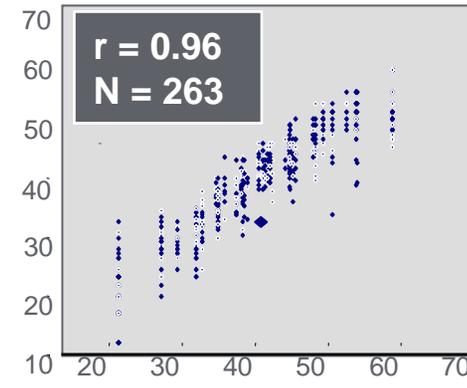
Do 5-item CAT Scores Agree with Criteria?



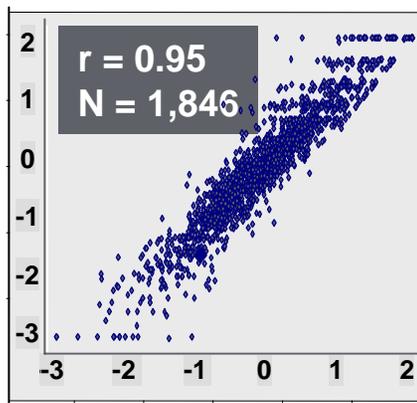
Mental Health



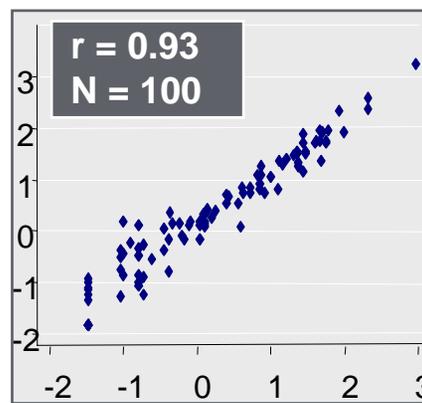
Headache Disability



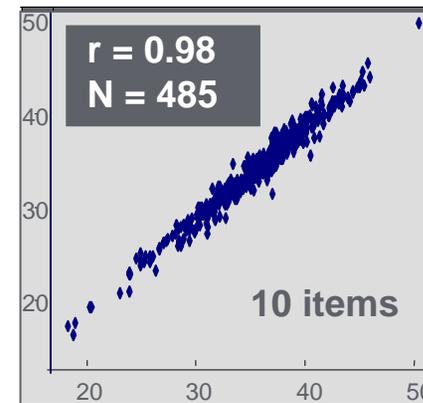
Pediatric Disability



Chronic Kidney Disease



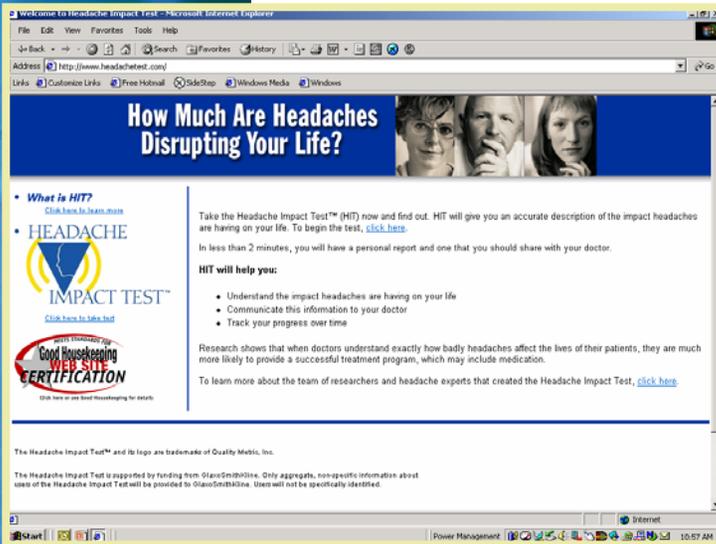
Diabetes



Post Acute Rehabilitation

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**



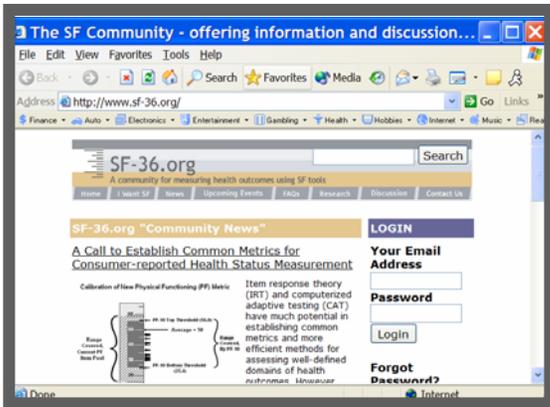
3rd Solution: The Internet

Reference: MS Bayliss, JE Dewey, R Cady et al., A Study of the Feasibility of Internet Administration of a computerized health survey: The Headache Impact Test (HIT), *Quality of Life Research*, 2003, 12: 953-961

New NIH Initiative:

Patient-Reported Outcomes Research Information System (PROMIS)

- NIH director's "NIH Roadmap: Re-Engineering the Clinical Research Enterprise"
- \$25M in NIH funding for a nationwide network to improve the monitoring of health outcomes
- Implement computerized adaptive test (CAT) assessment technology for administering questionnaires
- Target wide variety of chronic diseases.



We invite you to visit the following websites:

SF-36, SF-12 & SF-8 Health Surveys
www.sf-36.org

Headache Impact Test (HIT)
Asthma Control Test (ACT)
www.amlhealthy.com

Internet-based Health Assessments
www.qualitymetric.com
www.amlhealthy.com

Translations of Health Surveys
www.iqola.org

NIH PROMIS Initiative
www.nihroadmap.nih.gov/clinicalresearch/promis.asp