

Health-Related Quality of Life - 2006

John E. Ware, Jr., PhD

**QualityMetric Incorporated, Lincoln, RI
Health Assessment Lab, Waltham, MA
Tufts University School of Medicine**

**Introduction to the Principles and Practice of
Clinical Research
National Institutes of Health
Warren G. Magnuson Clinical Center
Building 10, Lipsett Amphitheater, Bethesda, MD
January 24, 2006**

Health Assessment Issues

- **Conceptual issues**
- **Standardizing metrics across applications**
- **Computerized dynamic health assessment**
- **Matching assessment methods to applications**

Quality of Life

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

Source: Campbell, 1981

What is Your Health?

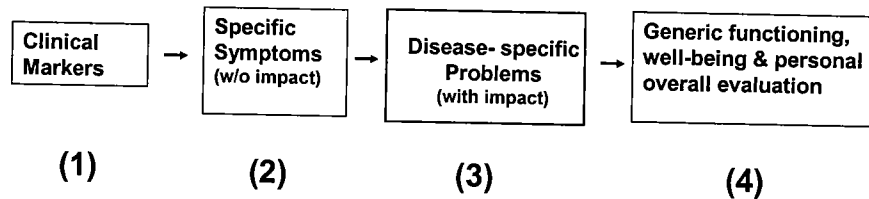
- Bodily structure & function
- What you are able to do – functioning
- How you feel – distress & well-being
- What you say it is – personal evaluation

Sources: Understanding Health Outcomes Educational Series

Background

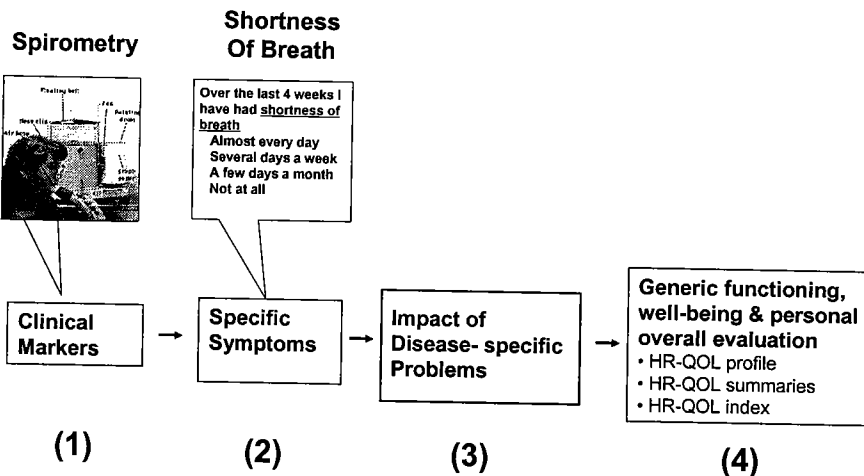
- Patient-reported outcomes (PROs) are very useful
- Standardization of concepts & metrics is enabling comparisons & interpretation
- Widespread use proves that more practical short forms will be adopted
- Traditional “static” short forms rarely meet clinical standards of precision
- Advances in assessment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

Continuum of Disease-Specific and Generic Health Measures



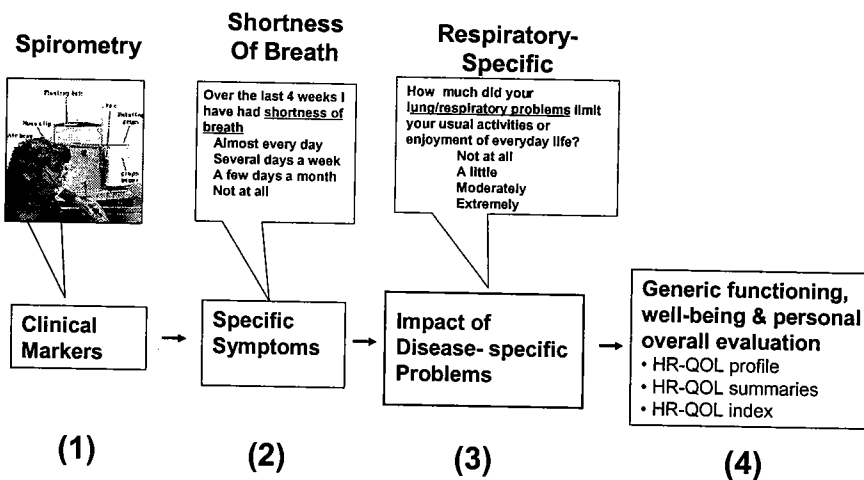
Adapted from: Wilson and Cleary, *JAMA*, 1995
Ware, *Annual Rev. Pub. Health*, 1995

Continuum of Disease-Specific and Generic Measures



References: Money et al., *J. Vasc Surg* 1998, 27(2):267-74
 Hiatt WR, *New Eng J Med*, 2001, 344(21), 160801621
 Regensteiner et al, *J Amer Geriat Soc*, 2002, 50, 1939-1946

Continuum of Disease-Specific and Generic Measures



References: Money et al., *J. Vasc Surg* 1998, 27(2):267-74
 Hiatt WR, *New Eng J Med*, 2001, 344(21), 160801621
 Regensteiner et al, *J Amer Geriat Soc*, 2002, 50, 1939-1946

Which is Most Valid for Purposes of Measuring HR-QOL?

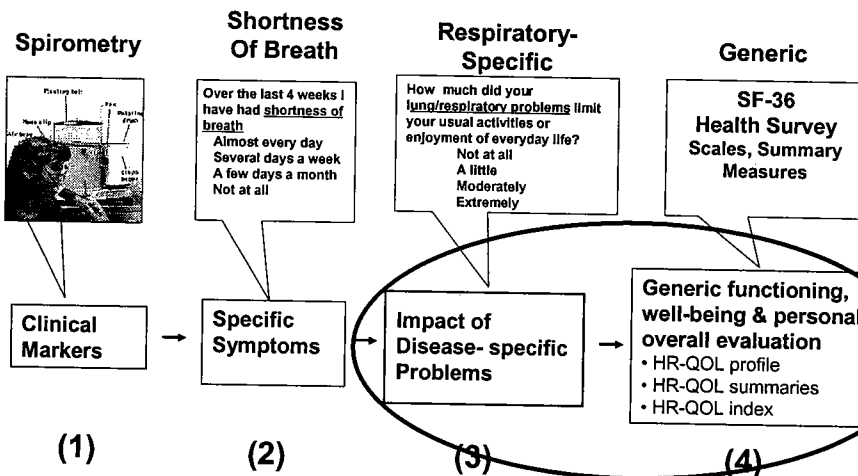
Over the last 4 weeks I have had shortness of breath
 Almost every day
 Several days a week
 A few days a month
 Only with lung/respiratory infections
 Not at all

(1) Frequency of breathlessness

In the past 4 weeks, how much did your lung/respiratory problems limit your usual activities or enjoyment of everyday life?
 Not at all
 A little
 Moderately
 Quite a lot
 Extremely

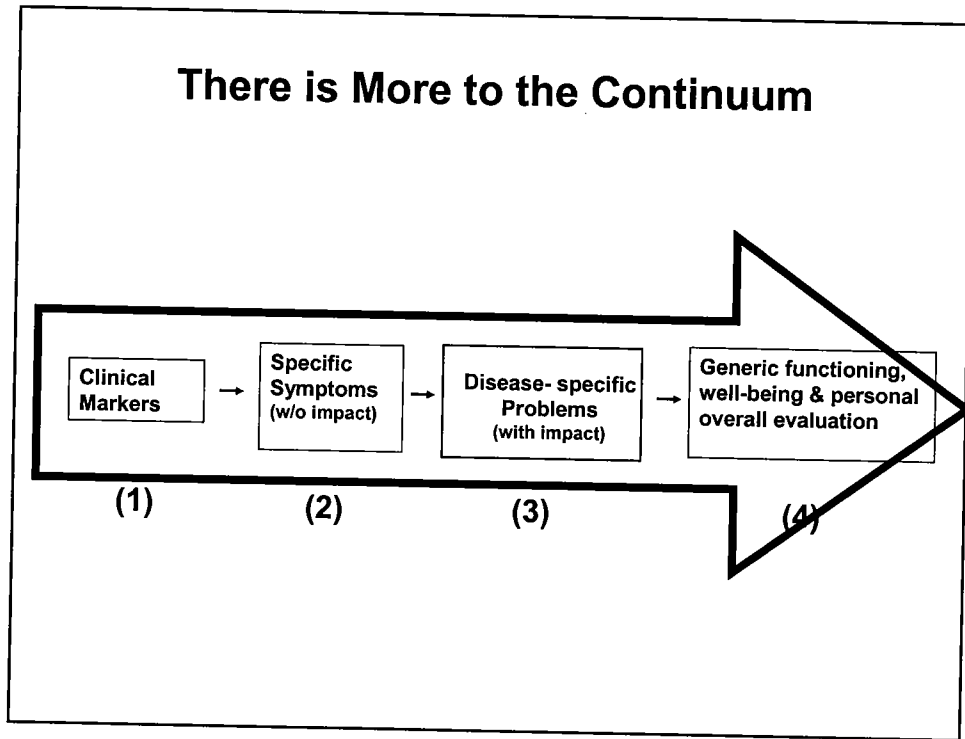
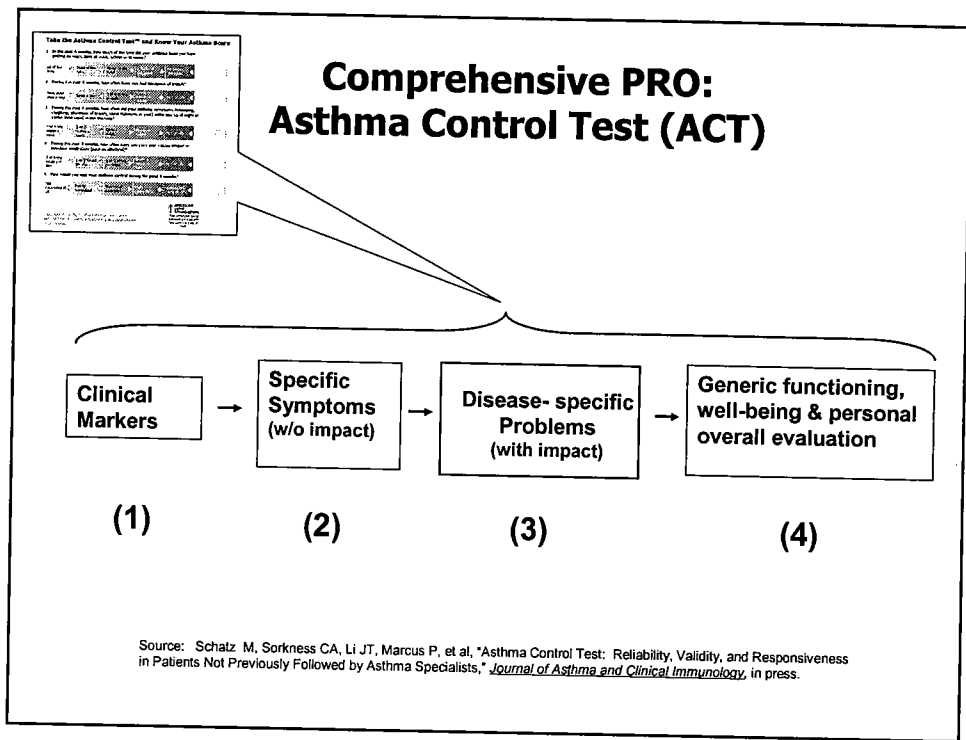
(2) Limitations in activities and enjoyment

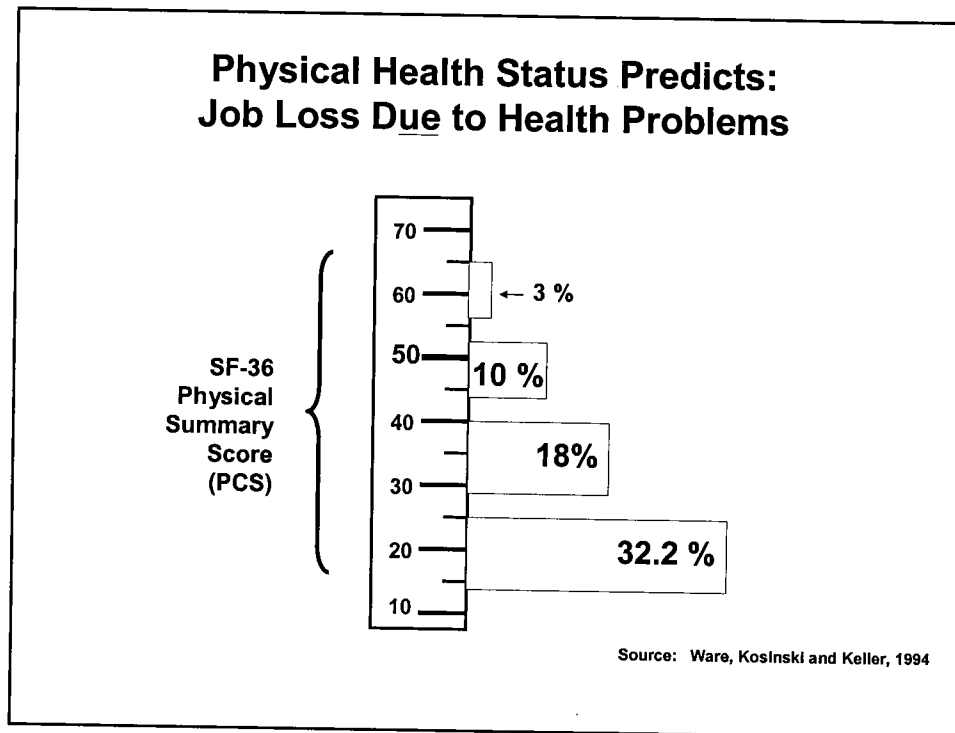
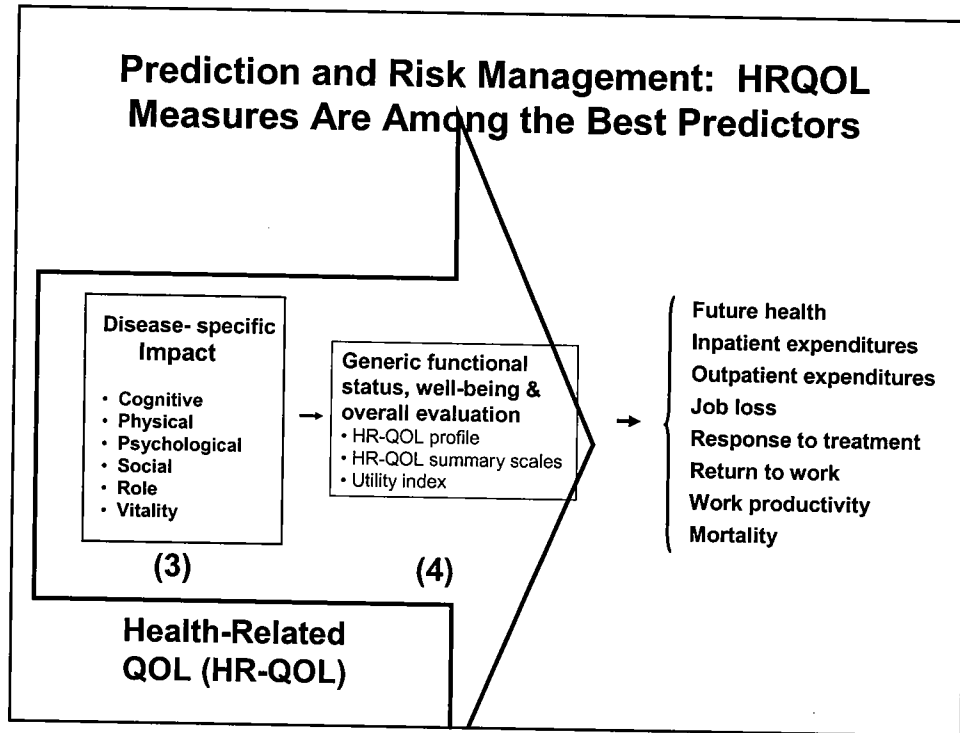
Continuum of Disease-Specific and Generic Measures



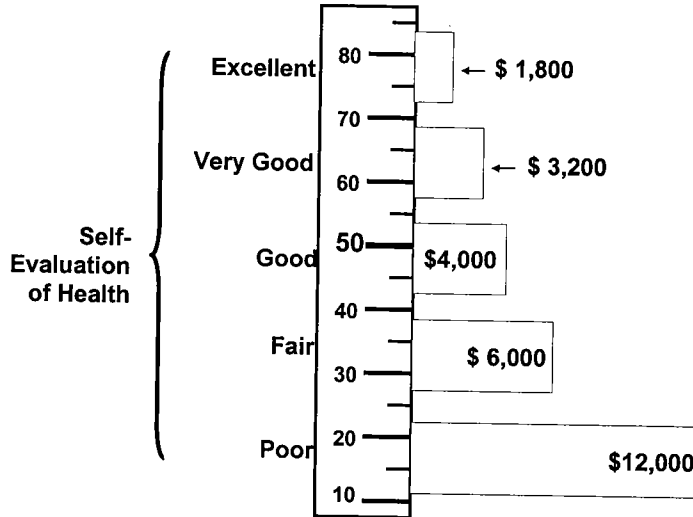
References: Money et al., *J. Vasc Surg* 1998, 27(2):267-74
 Hiatt WR, *New Eng J Med*, 2001, 344(21), 160801621
 Regensteiner et al, *J Amer Geriat Soc*, 2002, 50, 1939-1946

Health-related QOL



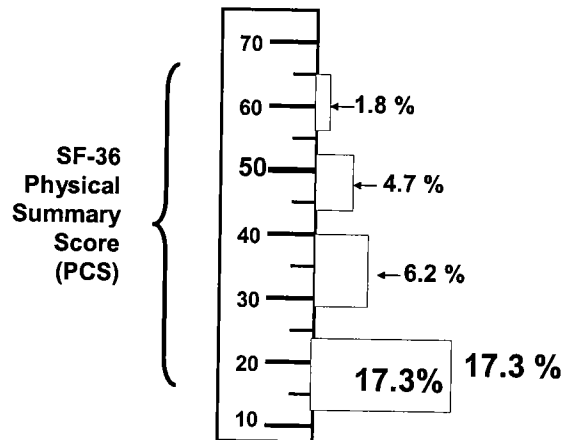


Self-Assessed Health Status Predicts: Annual Costs of Health Care



Source: Health Care Fin. Rev., 1997 vol 18 # 4

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



Source: Ware, Kosinski and Keller, 1994

What Do We Need for Clinical Research & Practice?

Physical activity limitations
Symptoms of psychological distress
Physical well-being
Life satisfaction
Emotional behavior
Role disability due to physical problems
Psychological well-being
General health perceptions
Physical mobility
Role disability due to emotional problems
Satisfaction with physical condition
Social activities with friends/relatives

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

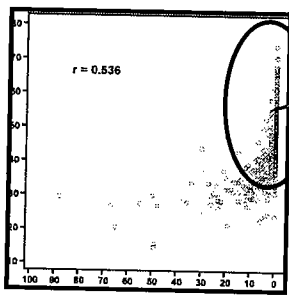
Summary of Information About Widely-Used General Health Surveys

Concepts and Characteristics	Psychometric										
	SIP	HIE	NHP	QLI	COOP	DUKE	MOS	QWB	EURO	HUI	SF-6D
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	•	•	•	•	•	•	•	•	•	•	•

Source: Adapted from Ware, 1995

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



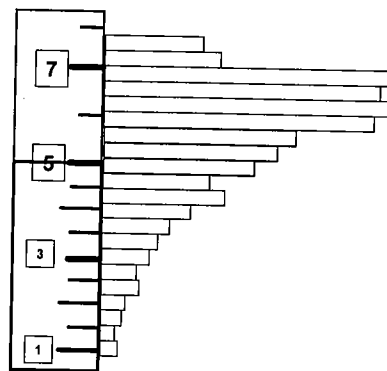
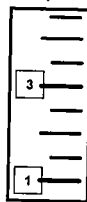
“Ceiling Effect”

What Do We Need for Clinical Research and Practice?

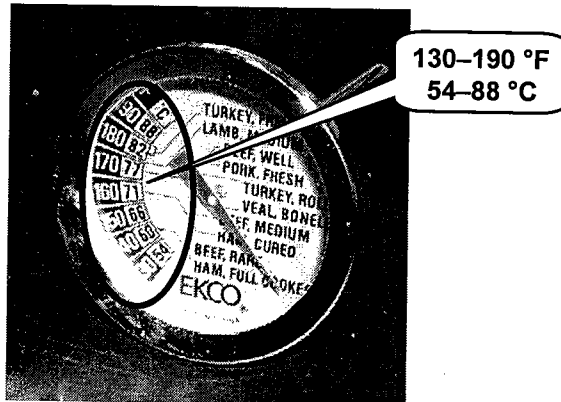
- Outcomes that matter to patients
- Practical measures
- Coverage of a wide range
- Greater precision
- Comparability of scores
- Ease of interpretation

Short-Form Surveys Do Not Provide a Solution

Measuring Too Low - Ceiling Effect

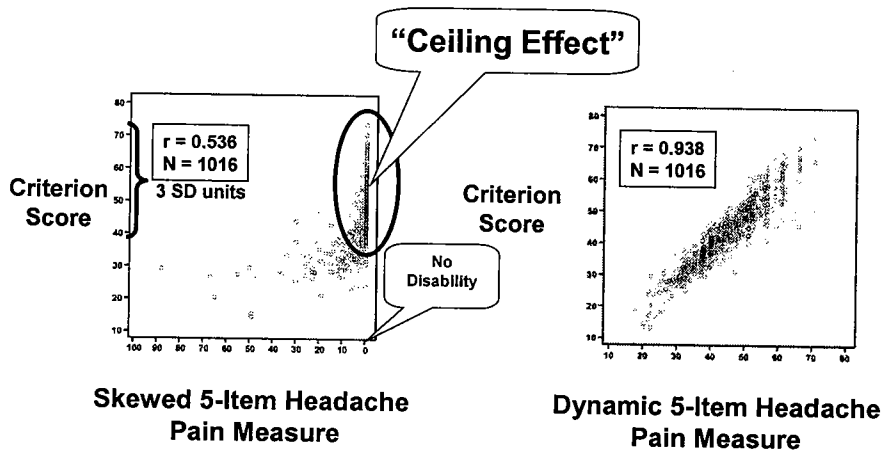


Some Thermometers Focus on a Very Narrow Range



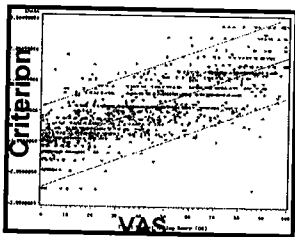
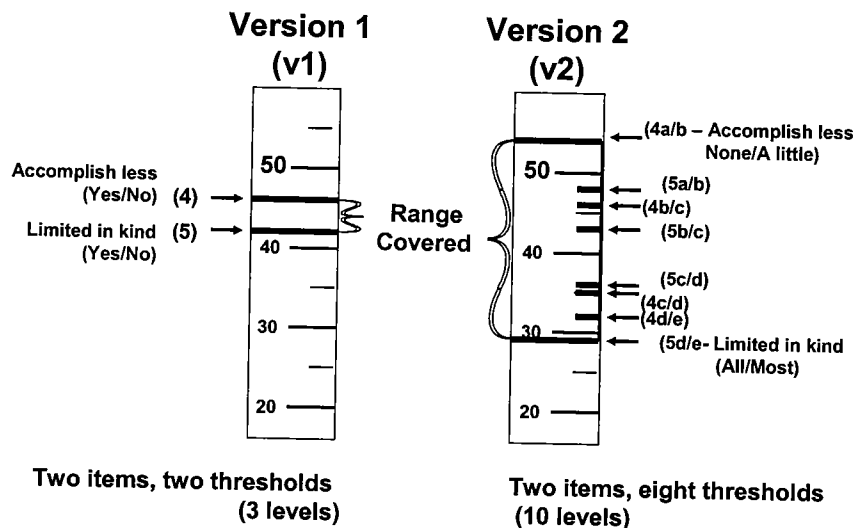
Cooking Thermometer

Another Solution is Computerized Dynamic Health Assessment



Ware JE, Jr, et al. *Med Care*. 2000;38:1173-82.

Comparison of Ranges Covered by SF-12 Role Functioning Items, v1 vs. v2



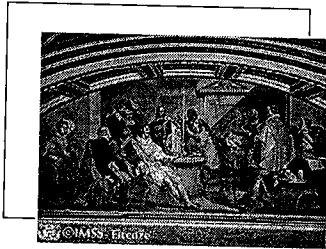
What Do We Need for Clinical Research and Practice?

- 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



Pain As Bad
As It Could Be

No Pain

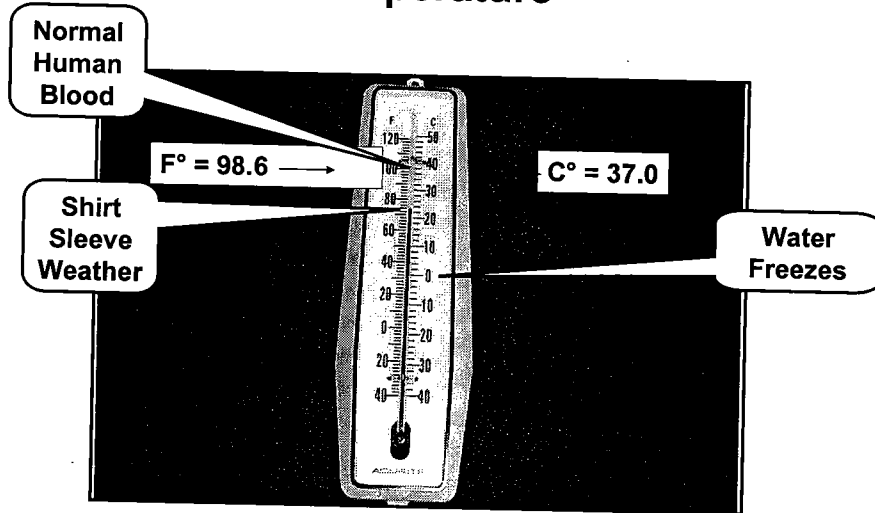
Visual Analogue
Scale (VAS)

What Do We Need for Clinical Research and Practice?

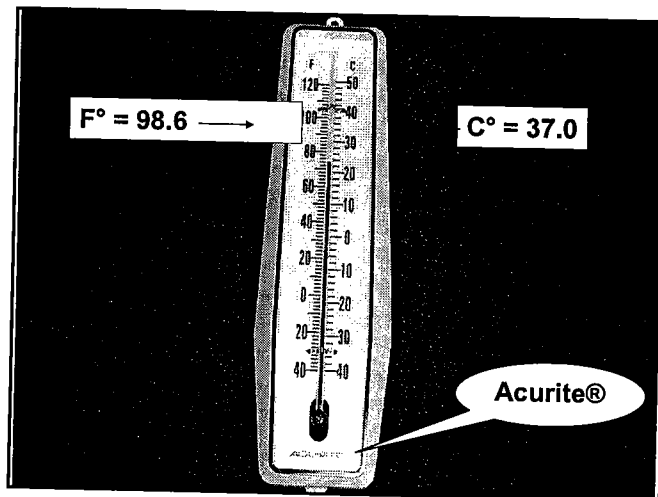
- Outcomes that matter to patients
- Practical measures
- Coverage of a wide range
- Greater precision
- Comparability of scores
- Ease of interpretation

Example: Cross-Calibrating Celsius and Fahrenheit

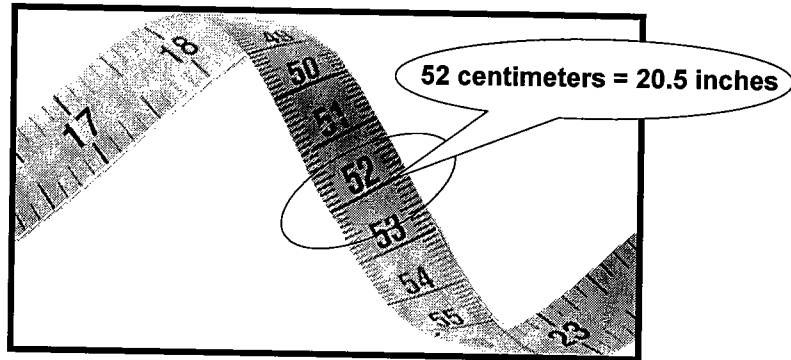
Temperature



Does Acurite® Own Temperature?



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



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

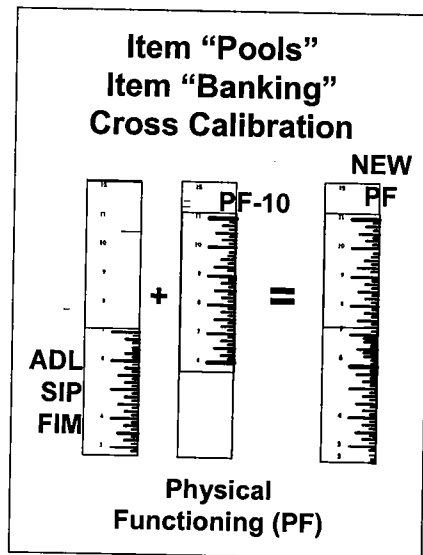
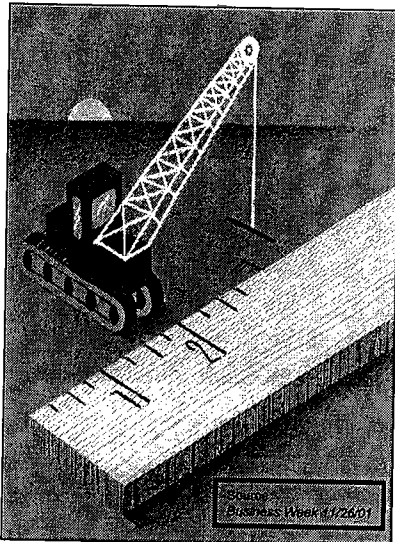
What Do We Need for Clinical Research and Practice?

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

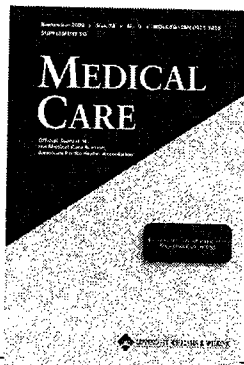
Solutions

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

First, Construct Better Metrics



Practical Implications of CAT in Health Assessment



MEDICAL CARE
Volume 44, Number 9, Supplement 9, pp B-73-B 62
©2009 Lippincott Williams & Wilkins, Inc.

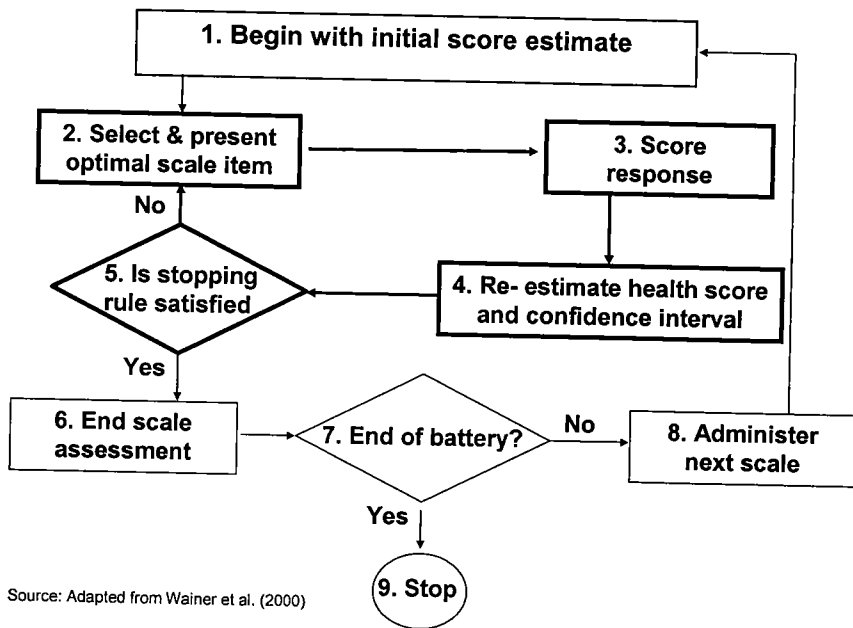
Practical Implications of Item Response Theory and Computerized Adaptive Testing A Brief Summary of Ongoing Studies of Widely Used Headache Impact Scales

JOHN E. WARE, JR, PhD,[†] JAKOB B. BOERNER, MD, PhD,[‡] AND MARK KOSINSKI, MA*

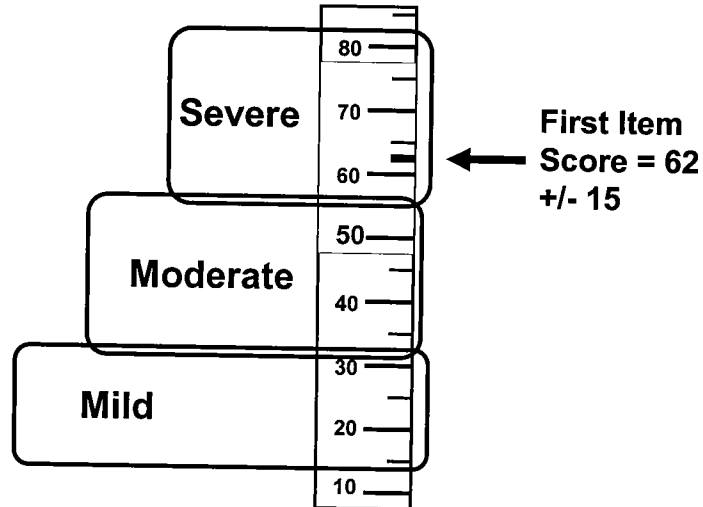
We have the potential to substantially advance the field of health status assessment by constructing and calibrating questionnaires based on item response theory (IRT) and administering them using computerized adaptive methods. This opportunity could

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

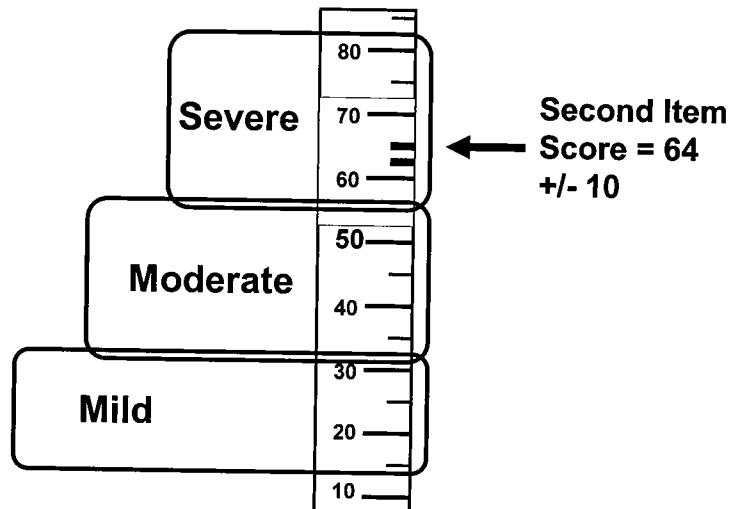
Logic of Computerized Adaptive Testing (CAT)



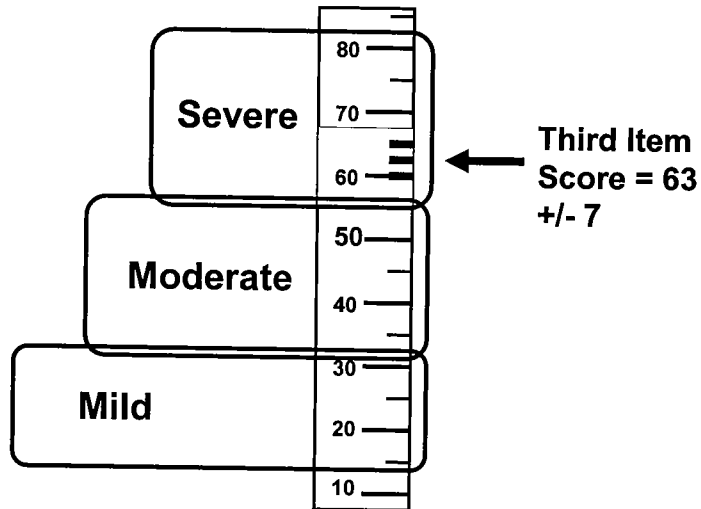
First Question Noisy Score Estimate (± 15)



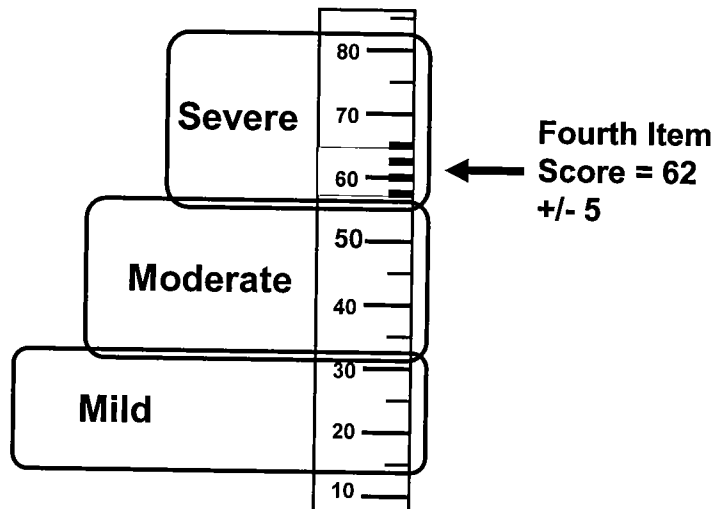
Second Question: Standard Error Reduced by One Third

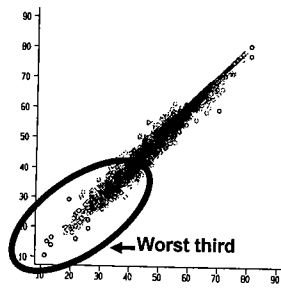


Third Question: Standard Error Cut in Half



Fourth Question: Standard Error Cut by Two Thirds





Do CAT Health Assessments Reduce Respondent Burden?

Results (N = 2,753):

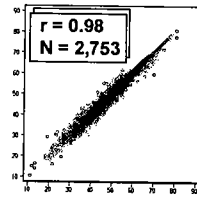
Among patients in poorest mental health (worst third):

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

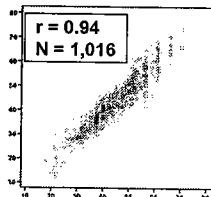
Acknowledgements

Functional Health CAT in Chronic Kidney Disease	NIDDK	Ware, Raczek
Functional Health CAT in Diabetes	NIDDK	Schwartz, Ware, Kosinski
CAT Assessment of Headache Impact	NINDS	Bjorner, Turner-Bowker, 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

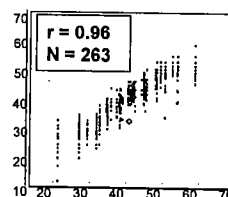
Do 5-item CAT Scores Agree with Criteria?



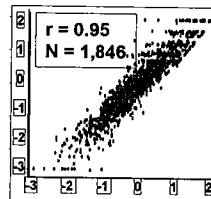
Mental Health



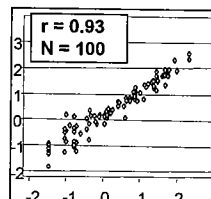
Headache Disability



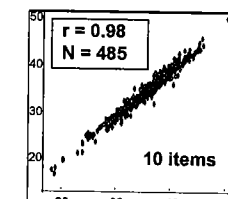
Pediatric Disability



Chronic Kidney Disease



Diabetes



Post Acute Rehabilitation



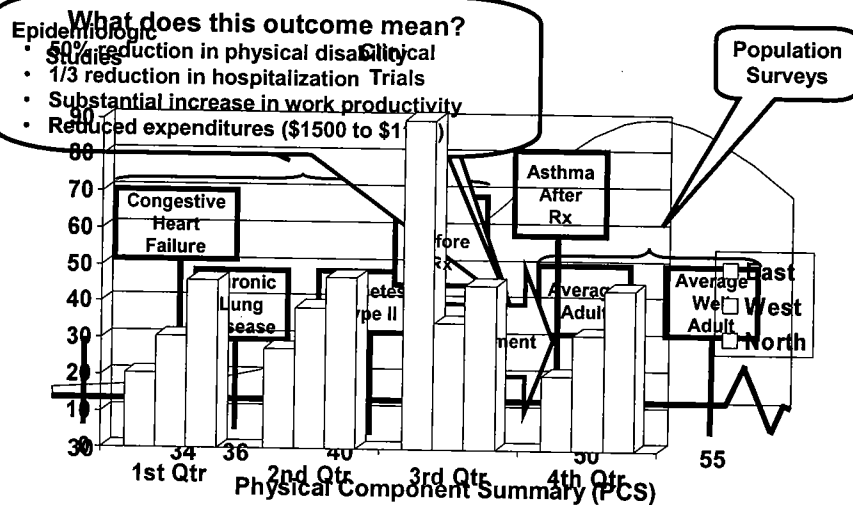
Articles Documenting the Development and Testing of the Headache Impact Test (HIT), *Quality of Life Research*, 2003, 12: 887-1012.

1. Bjorner et al., The Feasibility of Applying Item Response Theory to Measures of Migraine Impact: A Re-analysis of Three Clinical Studies
2. Kosinski et al., The Responsiveness of Headache Impact Scales Scored Using 'Classical' and 'Modern' Psychometric Methods: A Re-Analysis of Three Clinical Trials
3. Bjorner et al., Calibration of an Item Pool for Assessing the Burden of Headaches: An Application of Item Response Theory to the Headache Impact Test (HIT™)
4. Ware et al., Applications of Computerized Adaptive Testing (CAT) to the Assessment of Headache Impact
5. Bayliss et al., A Study of the Feasibility of Internet Administration of a Computerized Health Survey: The Headache Impact Test (HIT™)
6. Kosinski et al., A 6-item Short-Form Survey for Measuring Headache Impact: The HIT-6™
7. Gandek et al., Translating the Short-Form Headache Impact Test (HIT-6) in 27 Countries: Methodological and Conceptual Issues
8. Bjorner et al., Using Item Response Theory to Calibrate the Headache Impact Test (HIT™) to the Metric of Traditional Headache Scales
9. Turner-Bowker et al., Usefulness of the SF-8™ Health Survey for Comparing the Impact of Migraine and Other Conditions

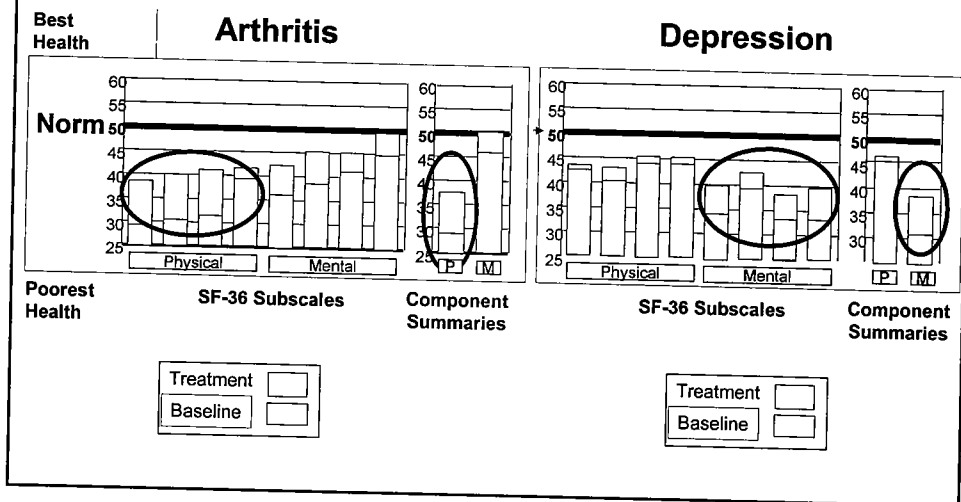
What are the Advantages of Dynamic Assessments?

- More accurate risk screening
- Reliable enough to monitor individual patient 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

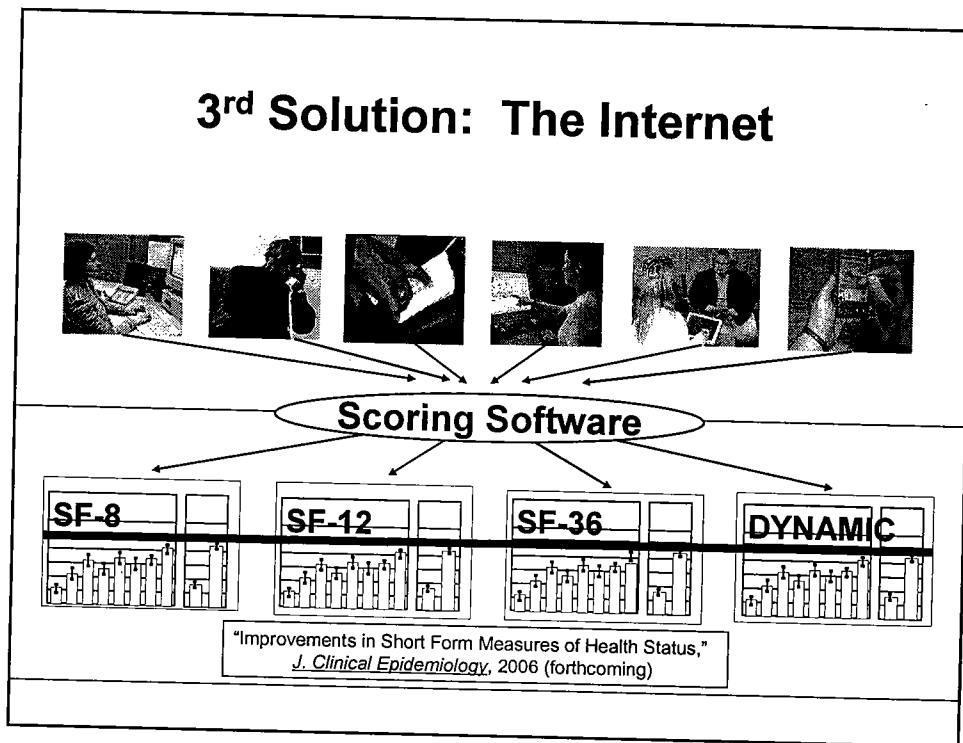
Putting it All Together: Standardization Makes Interpretation Possible



Chronic Disease Before & After Medication



3rd Solution: The Internet



BREATHEASIER, ACTNOW. - American Lung Association site - Microsoft Internet Explorer

- 130,000+ responses in first week
- 25,000+ 800# calls
- 105,000+ unique site visitors
- 75% completed ACT
- 63% were uncontrolled

Using the Internet to Administer an Asthma-Specific Screening Test (ACT)

www.lungusa.org

www.webmd.com

amlhealthy.com

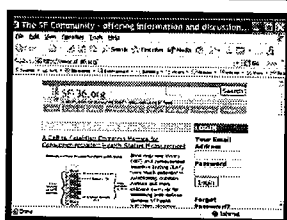
Using the Internet to Administer CAT SF-36 & Disease Impact Surveys

www.amlhealthy.com

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)
www.headachetest.com
www.amlhealthy.com

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

Translations of Health Surveys
www.igola.org

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