Patient-Reported Outcomes

Workshop

July 30-31, 2012

NATIONAL QUALITY FORUM

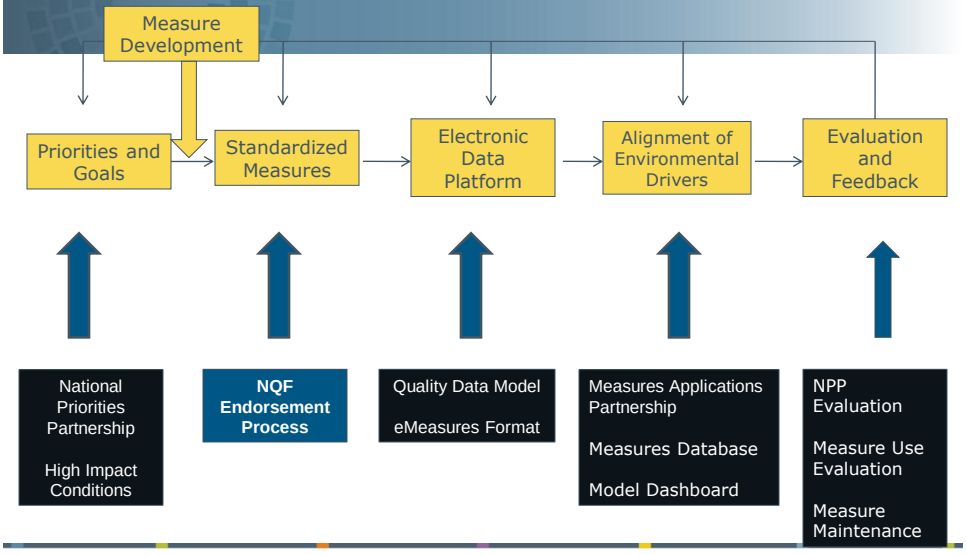
Project Scope and Activities

- Focused on methodological issues, not endorsement
- Under the guidance of an expert panel 2 workshops are planned to bring together the stakeholders necessary to facilitate the groundwork for the development, testing, endorsement, and implementation of PRO-based performance measures.
- Two commissioned papers will help inform next steps regarding: 1) selection of patient-level PROs for use in performance measures, and 2) the path to developing reliable and valid PRO-based performance measures eligible for NQF endorsement that can be used for accountability and to inform quality improvement
- Funded by HHS

Timeline

Call for nominations closed		4/2/12
Hold workshop #1		7/30-31/12
Expert Panel to discuss revision of first commissioned paper		8/21/12
Receive final version of first commissioned paper and prepare draft report of findings/recommendations		8/31/12
Hold workshop #2		9/11-12/12
Expert Panel review 2nd paper revisions & draft report for comment		10/11/12
Public/member comment period	open 10/23	11/21/12
Expert Panel to review comments received		12/3/12
CSAC and NQF Board review and approval		12/20/12

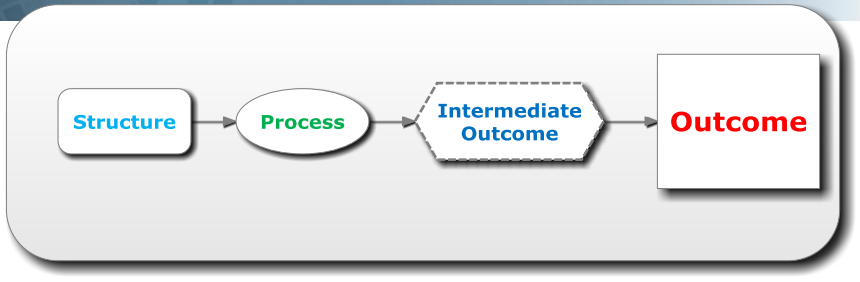
Quality Measurement Enterprise



Performance Measurement in Evolution

- Drive toward higher performance
- Measure disparities in all we do
- Shift toward composite measures
- Harmonize measures across sites and providers
- Measurement across longitudinal patient-focused episodes
  - Outcome measures (including PROs)
  - Process measures with direct evidence of impact on desired outcomes
  - Appropriateness measures
  - Cost/resource use measures coupled with quality measures, including overuse

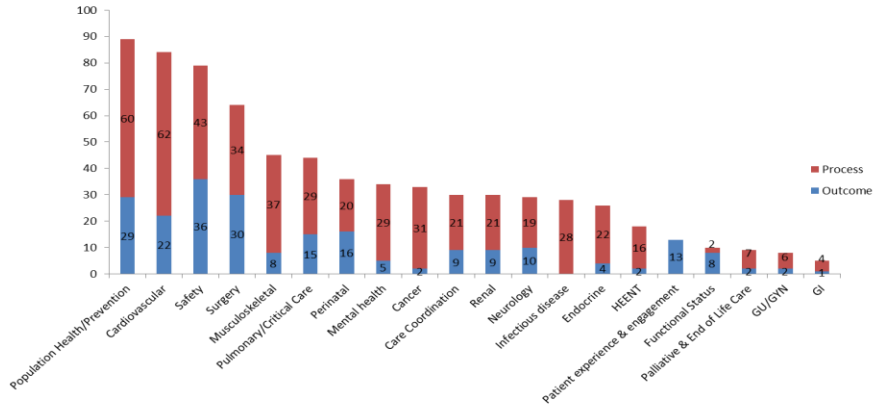
Evidence for the Measure Focus



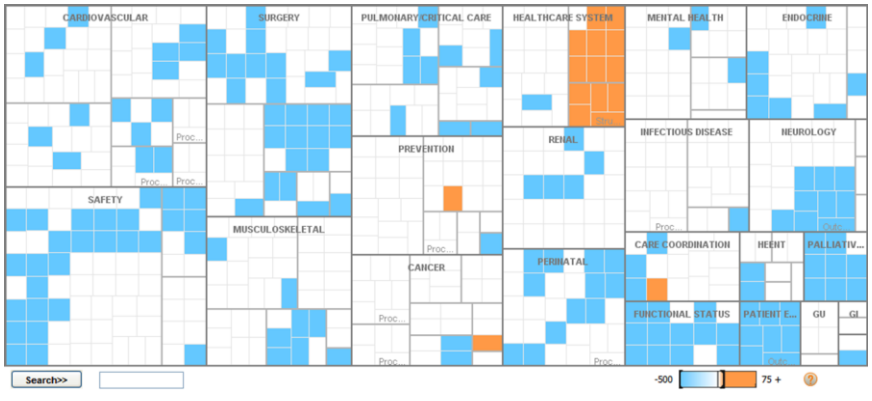
- Hierarchical preference for
  - Outcomes linked to evidence-based processes/structures
  - Outcomes of substantial importance with plausible process/structure relationships
  - Intermediate outcomes
  - Processes/structures

Most closely linked to outcomes

NQF Measure Portfolio: Process and Outcome Measures



NQF Measure Portfolio: Condition and Cross-Cutting Measures by Measure Types



Color legend: Process (white); Outcome (blue), Structure (orange)

Individual-Level PRO vs. Performance Measure

- **NQF does not endorse individual-level instruments or scales**
  - Although reliable and valid and useful in clinical practice or research, individual patient scores alone are not sufficient to determine performance and make conclusions about quality of a healthcare entity
  - Individual-level scores are the data that would be used in a performance measure
- **NQF endorses performance measures that result in a score for an accountable healthcare entity (and use data from all eligible patients)**
  - An endorsed performance measure must be standardized and precisely specified so specific instruments/scales and scoring must be identified

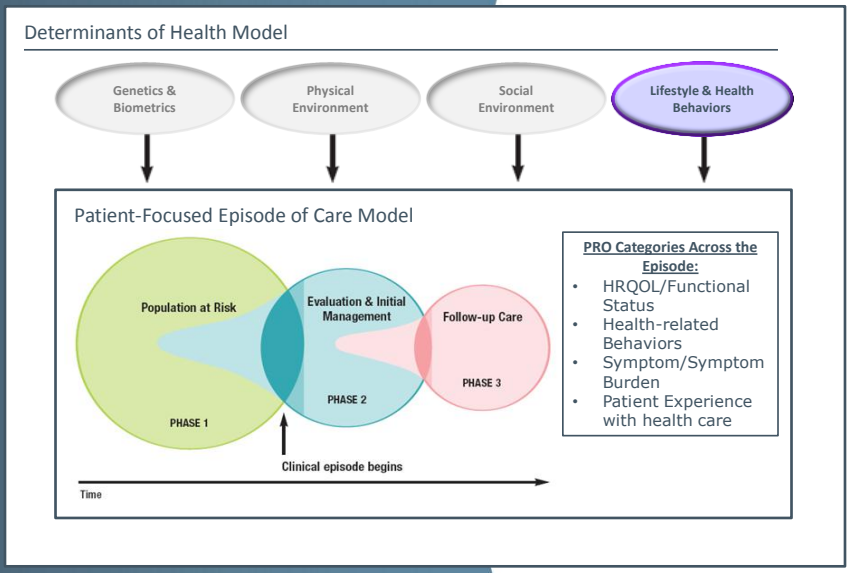
Examples: Endorsed PRO Performance Measures

- **Depression** (MN Community Measurement)
  - Depression Utilization of the Patient Health Questionnaire (PHQ-9) tool paired with:
    - Depression remission at six months
    - Depression remission at twelve months
- **Visual Function** (AAO)
  - Improvement in patient's visual function achieved within 90 days following cataract surgery
    - Improvement in visual function is defined by the quantitative scale used in the VF-14 survey instrument pre- and post-surgery.

Value Proposition

- Individual Level PRO
  - Inform care processes
  - Patient feedback and self monitoring
  - Shared decision-making
- Aggregate Level: Performance Measure
  - Quality improvement
  - Accountability (e.g., public reporting/transparency, payment)

Framing PROs Within Existing Conceptual Models



Today's Meeting Objectives

- Identify best practices and lessons learned from initiatives that have implemented individual-level PROs in performance measurement;
- Discuss the major methodological issues related to the selection, administration and use of individual-level PROs in performance measures;
- Discuss key considerations for inclusion of PROs into EHRs and policy implications;
- Identify the characteristics of individual-level PROs suitable for potential use in performance measures; and
- Identify an initial set of PROs most suitable for development and testing of performance measures.

Patient-reported outcomes in health care performance measurement: Issues related to selection and administration

NATIONAL QUALITY FORUM

David Cella, PhD
Department of Medical Social Sciences
Feinberg School of Medicine
Northwestern University

Co-authors

- Elizabeth A. Hahn
- Sally E. Jensen
- Zeeshan Butt
- Cindy J. Nowinski
- Nan Rothrock

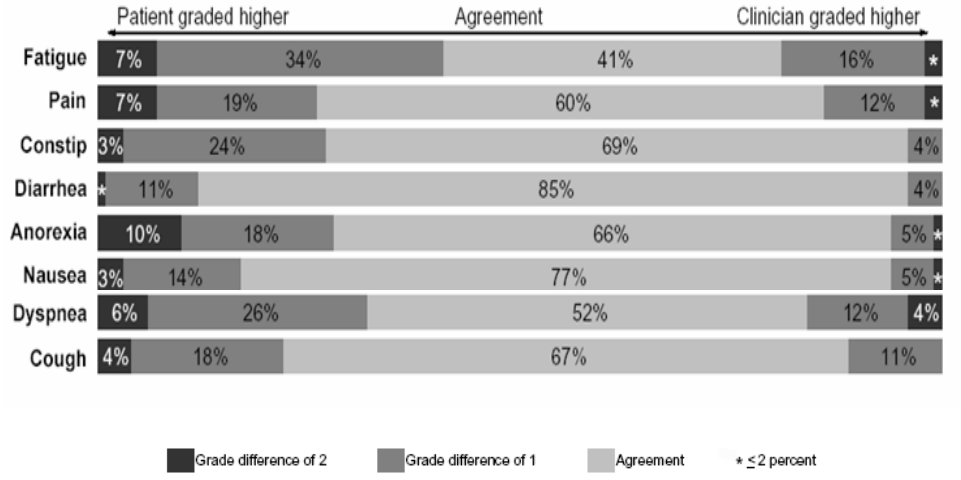
Acknowledging the Patient as an Authoritative Data Source

NATIONAL QUALITY FORUM

Why not just ask clinicians?

- Vast literature demonstrating that clinical providers do not accurately capture outcomes that are logically obtained by direct patient query

Comparison of Paired Observations



Basch, Lancet Oncol, 2006

Clinician AEs and Patient Reports: Lung Cancer

- **Uniscale** r = -0.06
- **Functional Assessment of Cancer Therapy Lung (FACT-L)** r = 0.10
- **Lung Cancer Symptom Scale (LCSS)** r = -0.03
- **Symptom Distress Scale (SDS)** r = -0.11

Clinician-reported AEs and PROs measure different aspects of the disease/treatment experience and are complementary

Bruner et al, 2008

Potential for PRO use in clinical care

- Assist clinical providers in care management
- Enhance clinical efficiency
- Improve patient-provider communication
- Identify patient needs in a timely manner
- Facilitate patient-centered care

However…

- Routine PRO assessment is not common in clinical practice

Patient Experience of Care: An integral component of patient-centered care

- Patient satisfaction

(example from FACIT-TS; www.facit.org)

Did your doctor seem to understand what was important to you?	No, not at all	Yes, but not as much as I wanted	Yes, almost as much as I wanted	Yes, and as much as I wanted

- Patient reports of their actual experiences

(example from CAHPS; www.ahrq.gov.cahps)

In the last 12 months, when you phoned this provider's office during regular office hours, how often did you get an answer to your medical question that same day?	Never	Sometimes	Usually	Always

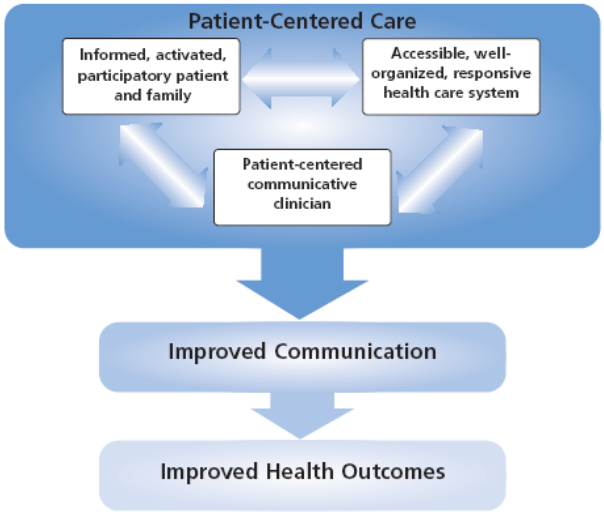
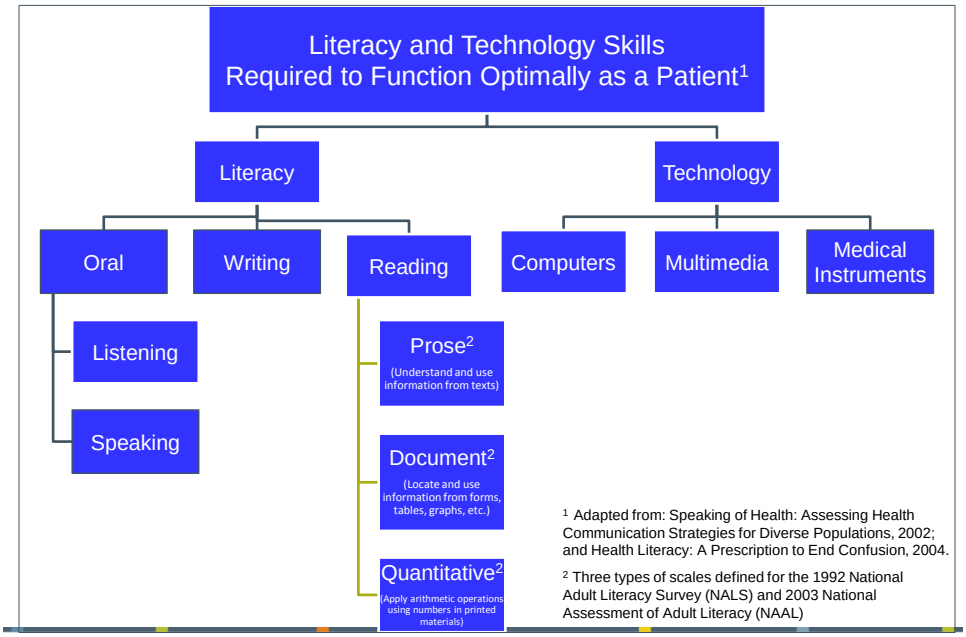


Figure 1.1 Based upon Epstein et al, Soc Sci Med, 2005, [from 2007 NCI/NIH Pub. #07-6225 "Patient-Centered Communication in Cancer Care"

Best practices to minimize self-report barriers

- Select appropriate method and mode of administration
  - Consider age, functional status, cognition as they relate to use of proxies and assistive devices.
- Universal design principles, quality translations and cultural adaptations, provide equivalent versions
- Flexibility in location (in-clinic, at home, at facilities)
  - requires access to the technology selected for the PRO
  - health information privacy must be protected
- Addressing *functional literacy* and *health literacy* are critical to delivering person-centered health care



Best practices to minimize barriers to self-reporting

- There are some circumstances in which it may be difficult or impossible to directly obtain PRO assessment by self-report.
- Proxy reporting can be useful:
  - for people with cognitive or communications deficits or severe disease burden
  - for people in the early stages of dementia who may fail to recognize the extent of their impairment
  - for young children

Discussion

Methodological Issues: Method of Administration, Collection & Response

NATIONAL QUALITY FORUM

Implications of Method/Mode of Administration and Response

- Decisions must be made related to data collection methods and the implications of those decisions on costs and errors in surveys (Groves, 2009)
  - What is the most appropriate method to choose for a particular question?
  - What is the impact of a particular method on survey errors and costs?
- Methods and modes differ along various dimensions:
  - Degree of interviewer involvement
  - Level of interaction with respondent
  - Channels of communication used
  - Degree of technology use

Source of Data and Methods/Modes of Survey Administration

- Self-report vs. proxy/observer ⟵ ***DATA SOURCE***
- Self-administration
  - paper-and-pencil
  - telephone
  - computer
- Interviewer-administration
  - paper-and-pencil
  - telephone
  - computer

MODE

METHOD

Implications of Data Source: Self Versus Proxy

- Proxy reporting is useful when difficult or impossible to obtain PROs directly
  - Allows broader inclusion and more representative range of patients
  - Minimizes missing data and increases the feasibility of longitudinal assessment
- Proxy reports may substitute for or complement patient assessment
  - May involve proxies assessing the patient as they think the patient would respond
  - May involve proxies providing their own perspective on the patient's status
- Evaluating agreement between patients and proxies
  - Greater agreement when rating observable functioning, activities of daily living, physical health, motor functioning and less agreement when rating social functioning, pain, cognitive status, psychological, emotional well-being
  - Magnitude of disagreement can be minimized
    - Disagreement may provide useful information (e.g., MCI → early dementia)

Implications of Mode of Administration

- Mode choices involve trade-offs and compromises
  - Consider the objectives of the assessment and the resources available
    - Self administration:
      - *Advantages: Cost; May yield more participant disclosure; Proceed at one's own pace*
      - *Disadvantages: Potential for missing data, Requires simple survey design*
    - Interviewer administration
      - *Advantages: Allows more complex survey design, Useful for patients with reading, writing, or vision difficulties*
      - *Disadvantages:: Cost; Potential for bias*
- Concern about the effects of mode on data quality
  - Reliability is high for both
  - Response effects favor self-administration but inconsistent

Implications of Method of Administration

- Paper-and-pencil
  - Advantages: Low start-up cost
  - Disadvantages: prone to data entry errors, data entry and scoring require more time, hard to incorporate into EHR
- Electronic:
  - Advantages: interactive, practical, increased comfort for socially undesirable behaviors, minimizes data entry errors, immediate scoring/feedback, easy to incorporate into EHR
  - Disadvantages: Up-front cost, potential discomfort with technology and accessibility
- Potential for differences between p & p versus electronic capture:
  - Impersonality of the method
  - Cognitive burden on patient
  - Control over the questionnaires
  - Communication style
- Increasing evidence of measurement equivalence between methods
  - As new methods are developed, it is critical to compare them to existing methods
- Across methods, patient privacy is always a concern

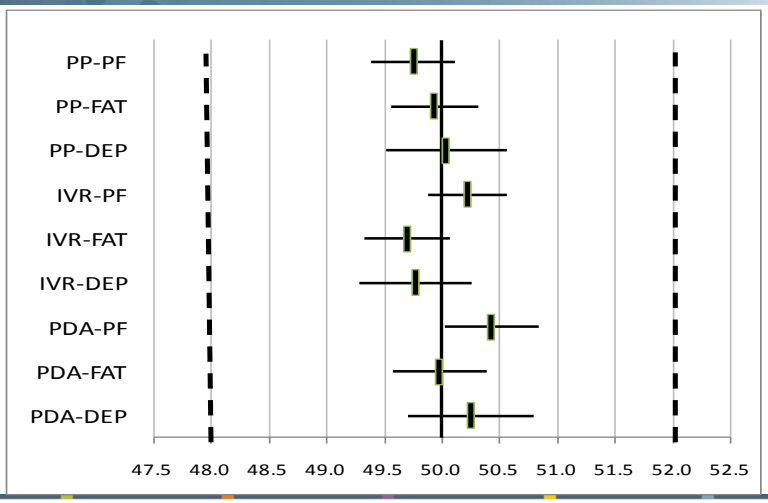
PROMIS Example: Method of Administration




No meaningful differences found between modes of administration

< 1.5 points on 100-point scale

Comparisons to PC Administration: MID > 2 points (0.2 SD)



People preferred the computer screen interface

Implications of Setting of Administration

- Clinic/Provider setting:
  - Strengths:
    - Real-time assessment
    - Easy to implement electronic administration
  - Limitations:
    - Impact on clinic flow
    - Interruptions result in missing data
    - Patient distraction/anxiety
    - Staff burden
- Home setting:
  - Strengths:
    - Minimizes impact on clinic flow
    - Minimizes staff burden
  - Limitations:
    - Accessibility
    - Health information privacy
    - Data security
    - Patient safety

Non-Response and Response Shift

- Bias may be introduced by missing data
  - Evaluate the amount, reasons and patterns of missing data
  - Apply statistical adjustments based on degree and pattern of missing data
  - Strategies to evaluate non-response bias:
    - Conduct an abbreviated follow-up survey with initial non-responders
    - Compare characteristics of respondents and non-respondents
    - Compare respondent data with comparable information from other sources
    - Compare early versus late respondents
- Patient adaptation and response shift over time can complicate the interpretation of PRO outcomes
  - Improvement may be unrelated to treatment effect
  - Consider monitoring for response shift or implementing control/comparator arms with longitudinal follow-up

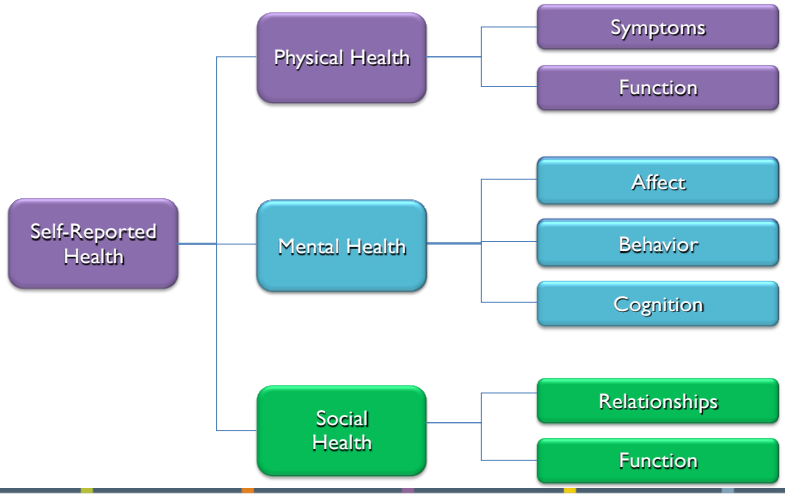
Implications of Test Theory Type

- Classical Test Theory (CTT): estimates the level of an attribute as the sum of responses to individual items
  - "Test-dependent": validity dependent upon all items to be completed
- Item Response Theory (IRT): "test-free": enables estimation of the latent trait using different items as long as their locations have been calibrated on the same scale as the patients' ability levels
- IRT enables customized assessment, including computer-adaptive testing (CAT) in which the questions are tailored to the individual patient
  - Questionnaires can be shorter
  - The scale scores can be estimated more precisely for any given test length
  - Patients do not need to complete the same set of items in every situation

PROMIS Demo

NATIONAL QUALITY FORUM

PROMIS Domain Framework



An item bank is a large collection of items measuring a single domain.

Any and all items can be used to provide a score for that domain.

The PROMIS Metric

T Score

Mean = 50

SD = 10

Referenced to the US General Population

Derived from Item Banks

Computerized Adaptive Testing (CAT)

- Dynamic testing averaging 6 items per domain

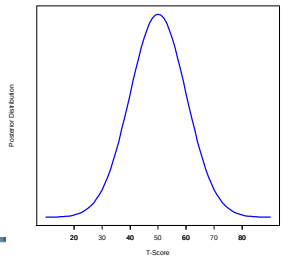
Fixed Length Forms

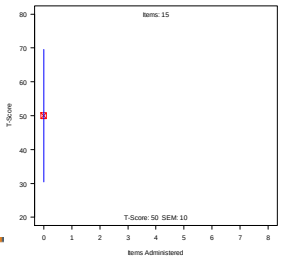
- By individual domain (8-10 items)
- By health profile (-29, -43, -57)

Global Health Index

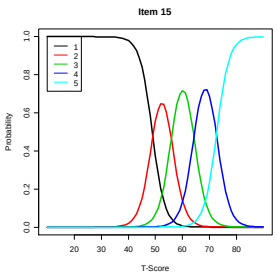
Beginning of CAT

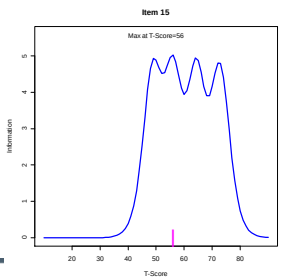
T-Score = **50** SE = **10**





Best Item-I felt depressed

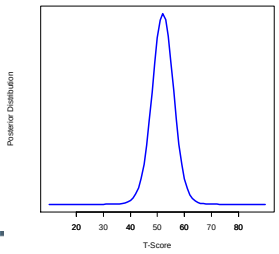


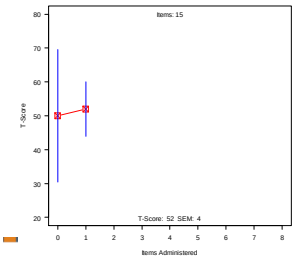


I felt depressed

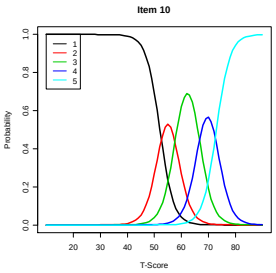
1. Never
2. Rarely
3. Sometimes
4. Often
5. Always

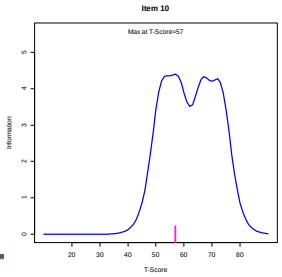
T-Score = **52** SE = **4**





Next Best Item-I felt like a failure

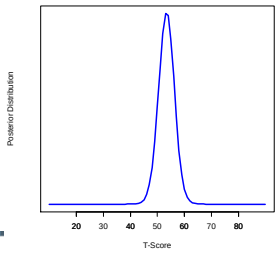


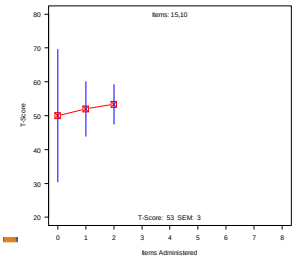


I felt like a failure

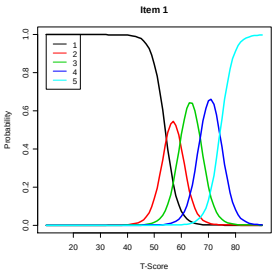
1. Never
2. Rarely
3. Sometimes
4. Often
5. Always

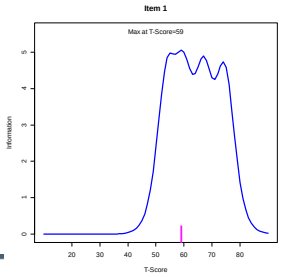
T-Score = **53** SE = **3**





Next Best Item-I felt worthless

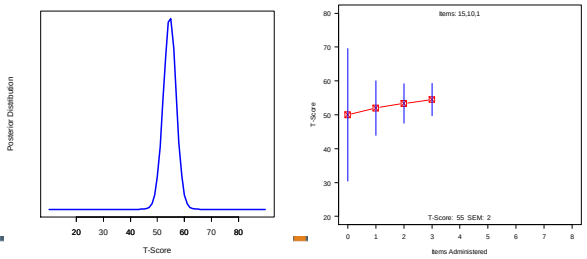




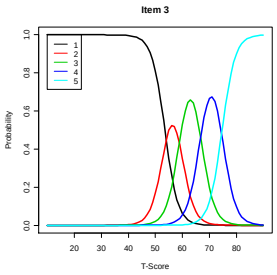
I felt worthless

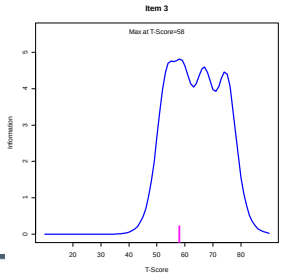
1. Never
2. Rarely
3. Sometimes
4. Often
5. Always

T-Score = **55** SE = **2**



Next Best Item-I felt helpless

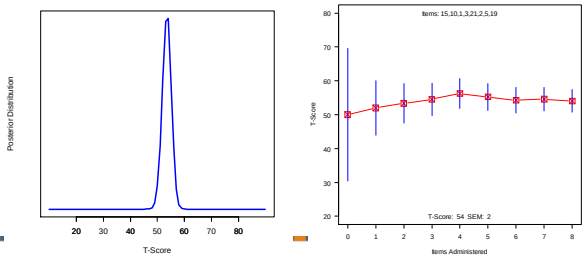




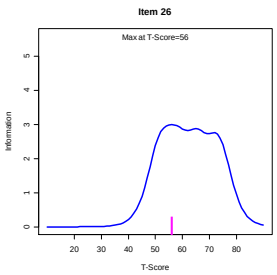
I felt unhappy

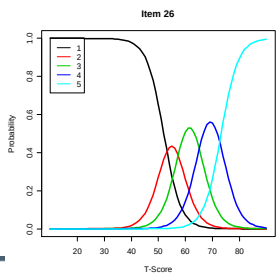
1. Never
2. Rarely
3. Sometimes
4. Often
5. Always

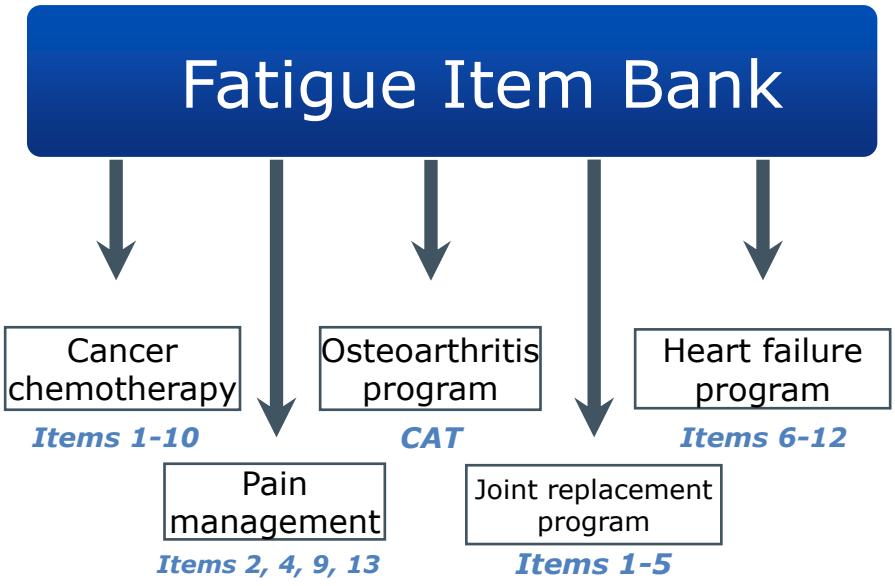
T-Score = **54** SE = **2**



Next Best Item-life was empty





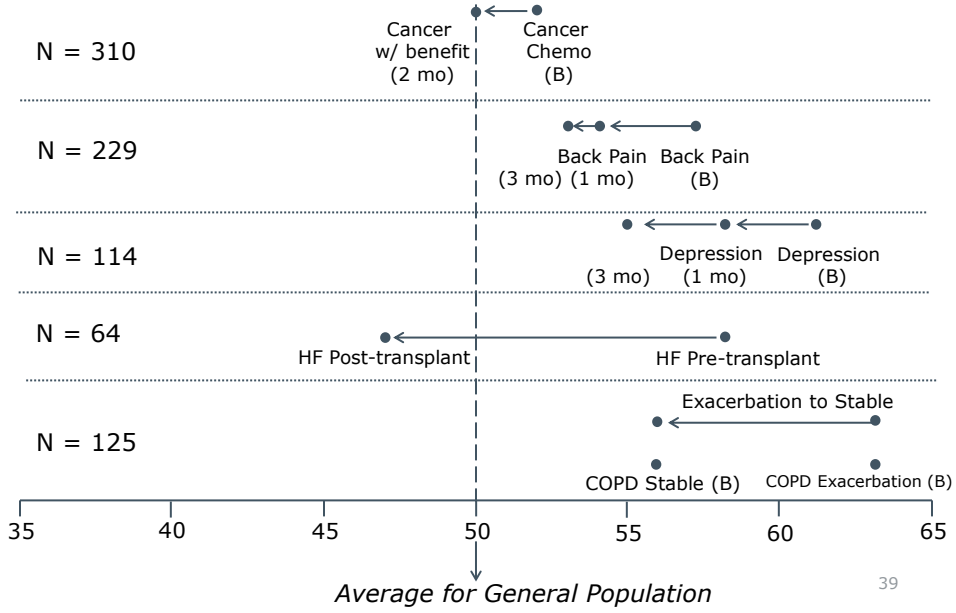


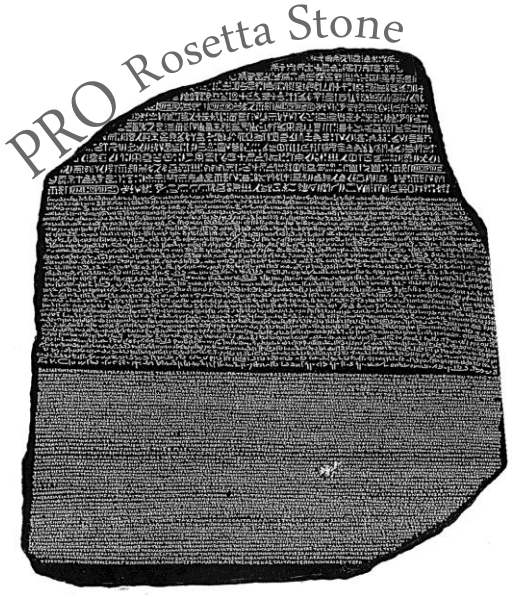
Same metric, same meaning

PROMIS Measures Tested in Six Conditions

Condition	Relevant Item Banks
COPD	Physical Function **Fatigue** Pain Social Role Satisfaction Emotional Distress (Depression, Anxiety, Anger)
Heart Failure	Physical Function **Fatigue** Social Role Satisfaction Depression
Low Back Pain	Pain (Interference and Behavior) Physical Function Depression **Fatigue** Sleep Disturbance
Depression	Emotional Distress (Depression, Anxiety, Anger) Sleep Disturbance **Fatigue** Physical Function Pain
Cancer	Pain **Fatigue** Emotional Distress (Depression, Anxiety) Physical Function

PROMIS Fatigue Across Five Clinical Conditions





PROsetta Stone Early Output

FACIT-F Score	PROMIS T-Score
52	27.8
51	32.8
50	35.9
49	38.4
48	40.3
47	42.0
46	43.4
45	44.8
44	45.8
43	46.9
42	47.9
41	48.8
40	49.8
39	50.5
38	51.3
37	52.1
36	52.8
35	53.6
34	54.3
33	55.0
32	55.7
31	56.3
30	57.0
29	57.6
28	58.3
27	58.9
26	59.5

FACIT-F Score	PROMIS T-Score
25	60.2
24	60.8
23	61.4
22	62.1
21	62.7
20	63.4
19	64.0
18	64.6
17	65.3
16	65.9
15	66.6
14	67.3
13	68.0
12	68.8
11	69.5
10	70.4
9	71.2
8	72.1
7	73.0
6	74.1
5	75.3
4	76.5
3	77.9
2	79.7
1	81.9
0	85.0

Smith et al, PM&R 2010: 2: 359-363

Discussion

Methodological Issues: Selecting Patient-level PROs

NATIONAL QUALITY FORUM

Selecting PROs for Performance Measurement

- To optimize decision-making in clinical care, PROs must be measured in a standardized way using questionnaires with known properties
  - Many guidance documents address attributes for PROs used in research
  - Little guidance regarding attributes for PROs used as performance measures

Selecting PROs for Performance Measurement

- Differences in selecting PROs as performance measures vs research
  - More similarities than differences
  - Importance of shorter instrument length
  - Higher stakes (consequences)
- Established PROs have more evidence than newer PROs …but newer PROs have better measurement properties
  - SF-36, SF-12, VR-36, VR-12 have been used most often
    - ➢ Limitation: Static measures
  - Future direction: IRT-based measures (e.g., PROMIS)

Characteristics of PROs Suitable for Use in Performance Measurement

- Review of recommended characteristics for PROs for use in performance measures
- *Example: Apply recommended PRO characteristics to the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC; Bellamy, 2008)*
  - *PRO for use in individuals with knee and hip osteoarthritis*
  - *24 items covering 1-14 days*
  - *5-point Likert-type and 100mm visual analog formats available*
  - *3 subscales:*
    - Pain (5 items)
    - Disability/Physical Function (17 items)
    - Joint stiffness (2 items)

Characteristics of PROs Suitable for Use in Performance Measurement

1. Conceptual and measurement model

- Documentation should include description of:
  - Concept(s) included and the intended population(s)
  - Organization of concept(s) into a measurement model
- Target PRO should be a high priority for the health care system
- Example: WOMAC
  - Factorial validity of the physical function and pain subscales has been inadequate (Pua et al., 2009)

Characteristics of PROs Suitable for Use in Performance Measurement

2. Reliability

- Internal consistency reliability should be:
  - ≥ 0.70 for group-level purposes
  - ≥ 0.90 for individual-level purposes
- Stability/Reproducibility depends upon the time window
- Example: WOMAC
  - Cronbach's alpha for the three subscales range from 0.86 to 0.98
  - Stability has been adequate for the pain and physical function subscales, but less adequate for the stiffness subscale

Characteristics of PROs Suitable for Use in Performance Measurement

3. Validity

- Evidence supporting:
  - Content validity
  - Construct validity
  - Criterion validity
- Limited number of PRO instruments have been validated for use in performance measurement
- Example: WOMAC
  - Development involved expert clinician input, and survey input from patients, as well as a review of existing measures
  - Patient ratings of satisfaction with arthroplasty correlate positively with WOMAC scores

Characteristics of PROs Suitable for Use in Performance Measurement

4. Responsiveness

- Evidence of changes in scores consistent with predefined hypotheses regarding changes in the target population
- Important for performance measurement because there is an expectation of consequences
- Responsiveness is necessary if results are to be actionable
- Example: WOMAC
  - Demonstrates adequate responsiveness and ability to detect change in response to clinical intervention

Characteristics of PROs Suitable for Use in Performance Measurement

5. Interpretability of scores

- Documentation should include:
  - What low and high scores represent
  - Representative mean and SD in the reference population
  - Guidance on estimating meaningful differences and change over time
- Performance measures:
  - If different PROs are used, important to establish a link or cross-walk
  - Application of criteria to determine clinically meaningful change
- Example: WOMAC
  - Availability of population-based, age- and gender-normative values
  - Availability of minimal clinically important improvement values
  - Can be translated into utilities for economic evaluations

Characteristics of PROs Suitable for Use in Performance Measurement

6. Burden

- Time, effort, and other demands on the respondent and the administrator
- Performance Measures:
  - PRO assessments should be as brief as possible
  - Reporting should be done in real-time
- Example: WOMAC
  - Short form available
  - Average time to complete mobile phone WOMAC = 4.8 minutes

Characteristics of PROs Suitable for Use in Performance Measurement

7. Alternative modes/methods of administration

- The use of multiple modes and methods can be useful for diverse populations
- Performance measures: Evidence of measurement equivalence necessary

- Example: WOMAC
  - ➢ Validated mobile phone and touchscreen platforms

Characteristics of PROs Suitable for Use in Performance Measurement

8. Cultural and language adaptations

- Performance measures: Mode, method and question wording must yield equivalent estimates of PRO measures

- Example: WOMAC
  - ➢ Available in over 65 languages

Characteristics of PROs Suitable for Use in Performance Measurement

9. Electronic health records

- Performance measures: Critical features:
  - Interoperability
  - Automated, real-time measurement and reporting
  - Sophisticated analytic capabilities

- Example: WOMAC
  - Electronic data capture may allow for integration within EHR

Discussion

Incorporating PROs into Electronic Health Records & Personal Health Records

NATIONAL QUALITY FORUM

E-health:
"Health-related Internet applications that deliver a range of content, connectivity and clinical care"

- health information
- online formularies and prescription refills
- appointment scheduling and test results
- advance care planning and health care proxy designation

➢ e-health applications tend to focus on the needs of health care providers and organizations

➢ there is little evidence regarding whether the services offered are those that patients desire

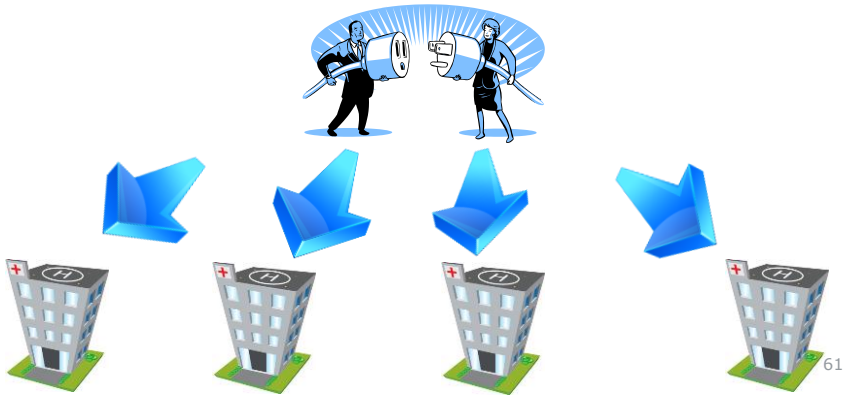
Integrating PROs into EHRs & PHRs

- PROs will likely constitute an important aspect of future stages of "meaningful use" of EHRs
- Critical features:
  - Interoperability and widespread health information exchange
  - Automated, real-time quality and cost measurement
  - Sophisticated analytic capacities

Integrating PROs into EHRs & PHRs

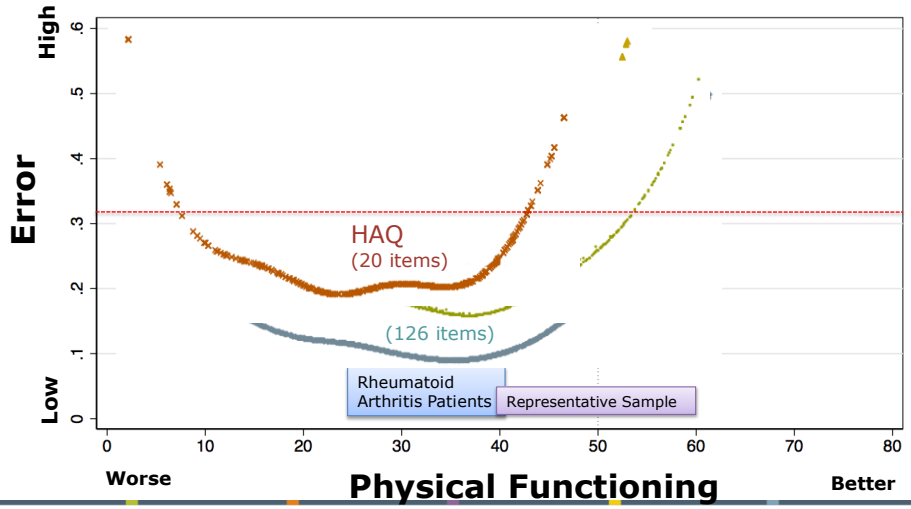
- Important issues:
  - Patient perspective:
    - Patients want to be involved as a partner in the flow of information
  - Clinical buy-in
  - Compatibility with clinical flow
  - Meaningful use
  - Patient privacy:
    - physical transfer of the paper-based PRO measure from patient to provider
    - electronic transfer of data or unauthorized access to patient-reported data
- Key design principles:
  - Fitting PRO measures into flow of care
  - Designing the system with stakeholder engagement
  - Merging PRO data with other types of data
  - Engaging in continuous improvement of the systems based on user experience and new technology

SNOMED and LOINC codes

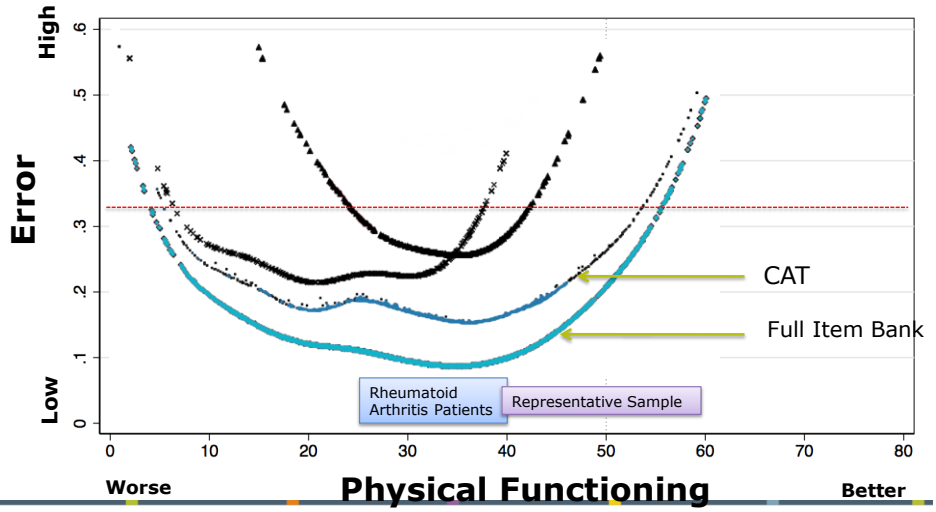


Discussion

Reliability/Precision of PROMIS
(Physical Function)



Relative Precision of PROs
(Physical Function)



Lessons from the field
Early experience with PROs

Partners HealthCare, Inc.
Boston, MA

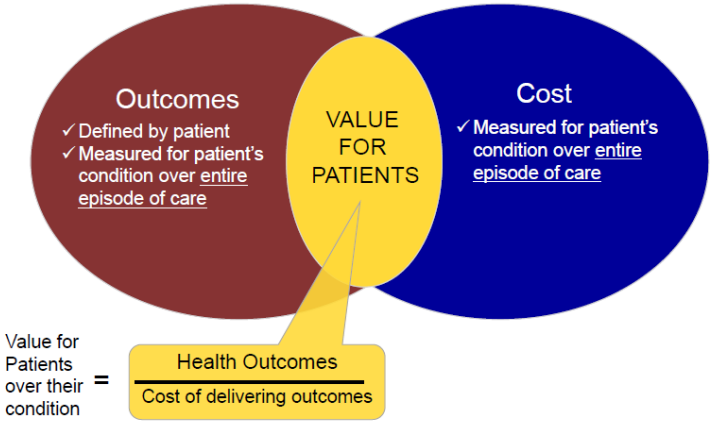
Elizabeth Mort, MD, MPH
Senior Medical Director PHS
VP Quality & Safety MGH

NATIONAL QUALITY FORUM

Partners Strategic Plan: Care Redesign

- October 2010 launched Strategic Plan at Partners
- Care Redesign
  - Primary care, population health
  - Condition specific care
    - CABG
    - Stroke
    - Colectomy for Colon Cancer
    - AMI
    - Diabetes

Key Guiding Principle behind Care Redesign



First wave of conditions: CABG & AVR and Diabetes

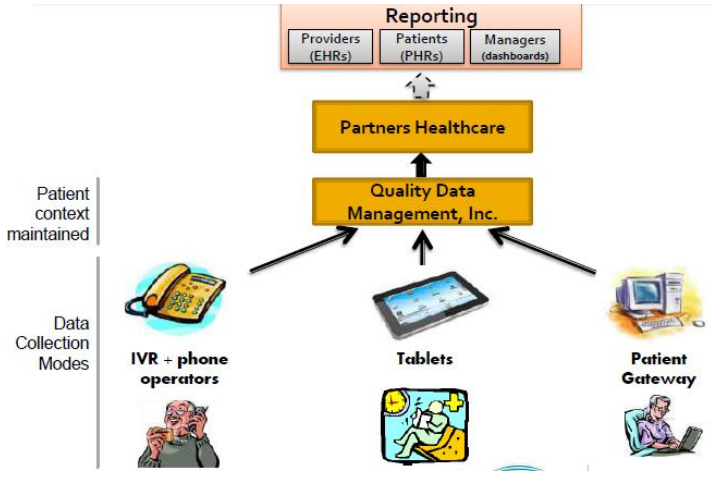
Goals: electronic, validated instruments, short, align

CABG	Pre-Procedure	Post-Procedure
FUNCTIONAL STATUS (PROMIS-10) (General, mental, social, physical, anxiety, fatigue, pain)	10	10
SYMPTOMS LEVEL (chest pain, shortness of breath)	6	6
PERCEIVED HEALTH BENEFITS (Perception of procedures' success, and physical/emotional improvement compared to yr ago)	-	4
HEALTH UTILITY (Health state from 1-100)	1	1
Total Questions:	17	21

Diabetes	Continuous
FUNCTIONAL STATUS (PROMIS-10) (General, mental, social, physical, anxiety, fatigue, pain)	10
ANXIETY (PROMIS) (worries, ability to focus, fearfulness)	4
BURDEN OF DIABETES (quality of life in light of disease, ability to cope w/ disease)	1
HEALTH UTILITY (Health state from 1-100)	1
Total Questions:	21

CABG: Pre-op Post-op 3,6,12 months | **Diabetes: baseline and every six months**

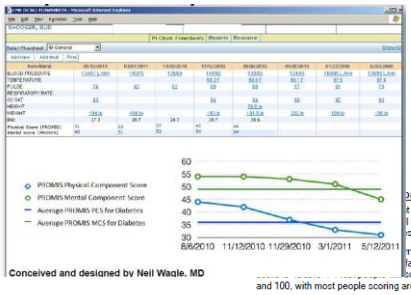
Collaborated with a vendor for data collection



We have been in the field since March 2012

- ***264*** **questionnaires completed since March 2012 (122 for cardiac surgery, CABG and AVR, and 142 for Diabetes).**
- **~*56*% of the patients choosing a method of follow up selected the patient portal/internet option (not Interactive Voice Response).**

Developing reporting formats and mechanisms, using electronic medical record and patient portal

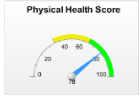


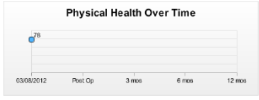
Conceived and designed by Neil Wagle, MD

bass surgery

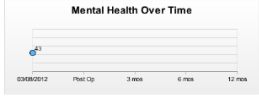
t Reported Outcomes Survey. From the answers you gave on the survey, we are able to give you a score in 3
l be compared to scores obtained from a national population that included young to old adults and also healthy
s

ry out various
fatigue. Your
score between 30
and 100, with most people scoring around 50.





Mental Health is your self-report of quality of life, your mood, and emotional feelings of anxiety, depression or irritability. Your score is <Score>



Early feedback from patients, staff, physicians

The patient experience:
- **Patients say their doctors "should be asking these questions"**
- **Patients comment that the tablets are fun to use and "very user-friendly"**
- **Patients are willing to answer these questions at home**

The staff experience:
- **Practice Administrators have created unique workflow plans for their clinics to best incorporate PROMs**
- **MA's and nurses generally understand the importance and provide guidance to the patients through the process**

The physician experience (preliminary):
- Variable response – from champions to critics
- Concern about fitting these data elements into the clinical encounter
- Concern about critical results and timely intervention

Lessons learned from planning and evaluation

- **Integrate this data collection with system-wide focus on improving value**
- **Significant up-front investment in research, interviews, patient focus groups, change management**
- **Establish senior executive, system wide clinical champions, local physician leadership, operational champions**
- **Careful attention to each clinic's unique workflow and organizational culture**
- **Engage providers in the design of reporting tools**

Promise of PROs in Improving Patient Outcomes: Lessons from the Dartmouth Spine Center

Eugene C. Nelson, DSc, MPH
The Dartmouth Institute
Dartmouth-Hitchcock Health

NATIONAL QUALITY FORUM

To learn more see: Using patient-reported information to improve health outcomes and health care value. Nelson, Hvitfeldt, Reid, et. al. Technical Paper, The Dartmouth Institute, June 2012.
http://tdi.dartmouth.edu/initiatives/engage/payment-models-delivery-system-reform/measurement

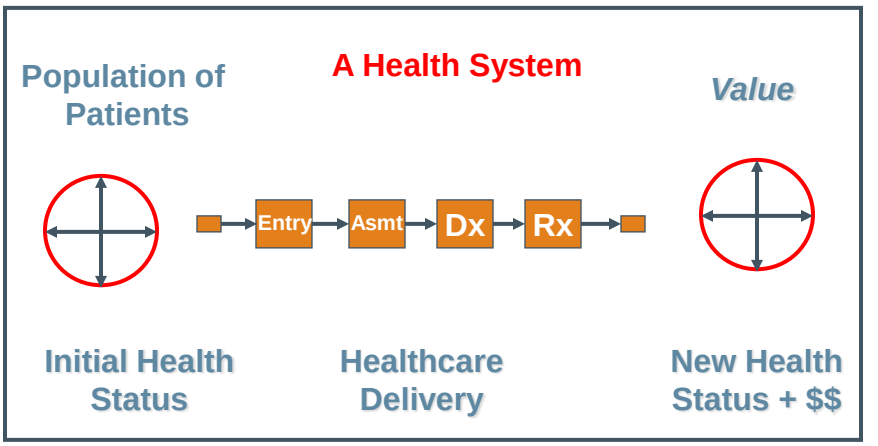
Feed Forward PROs to Improve Outcomes and Health Care Value: Dartmouth Spine Center Case

Q: How is a kilowatt hour of electricity like a day in the hospital?
A: Nobody wants either.
We want
- Cold beer & hot showers
- Better outcomes, better care & lower costs

- End use, least cost
- Value for money
  - *Amory Lovins*

Skating to where the puck is going to be … person-centered high value care

What is health care value?



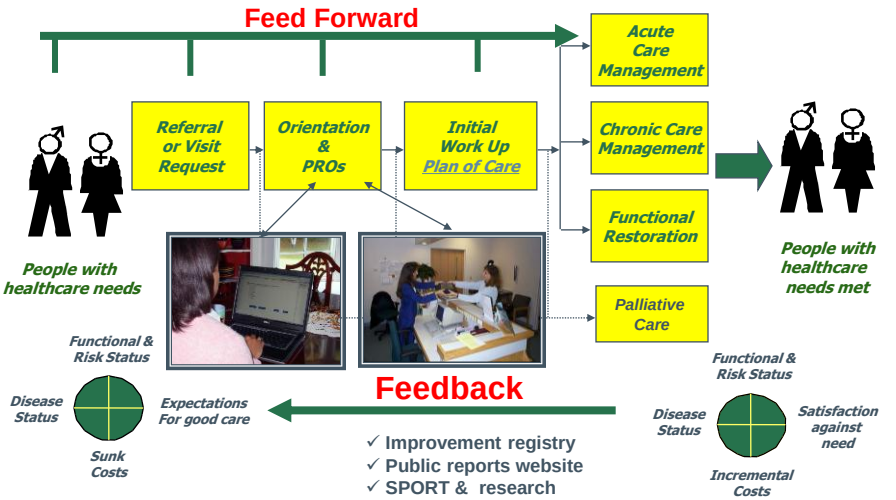
Value = Health outcomes (disease + risk + function) / costs over time

Dartmouth Spine Center

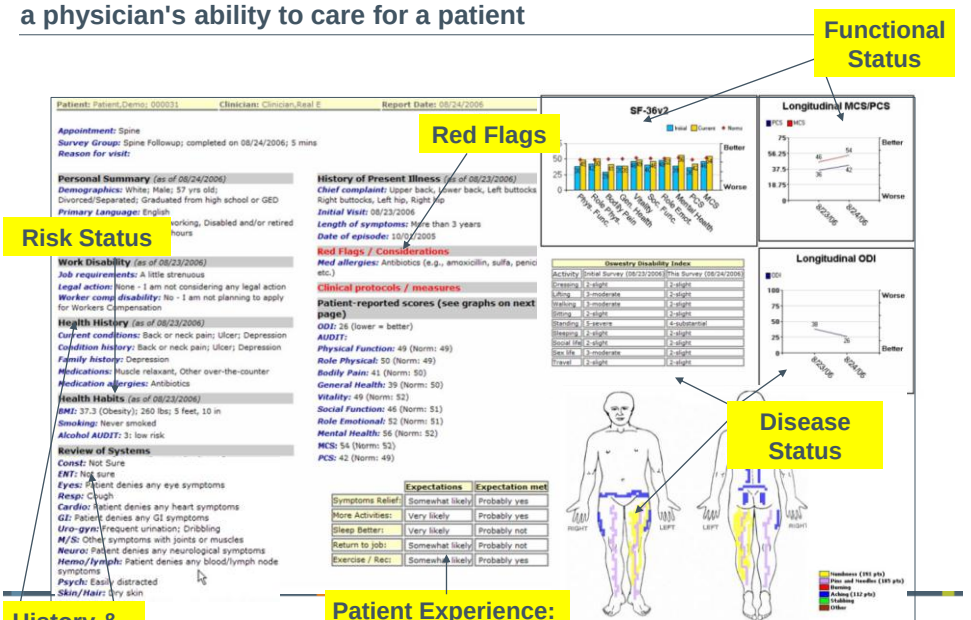
- Started in 1998 by Jim Weinstein
- Innovative interdisciplinary clinical microsystem … 1 stop shopping
- "Back to work back to play 1 back at a time." … patient-centered
- Better care in real time & better research over time
- "I can't be a good doctor if I don't have PROMs"

Dr. Jim Weinstein, orthopod & D-H President

Spine Center: Feed forward (& feedback) system, featuring PROs for engaging patient, shared decision making & making care plan, coordinating care, improving care, measuring, researching & paying for health care value



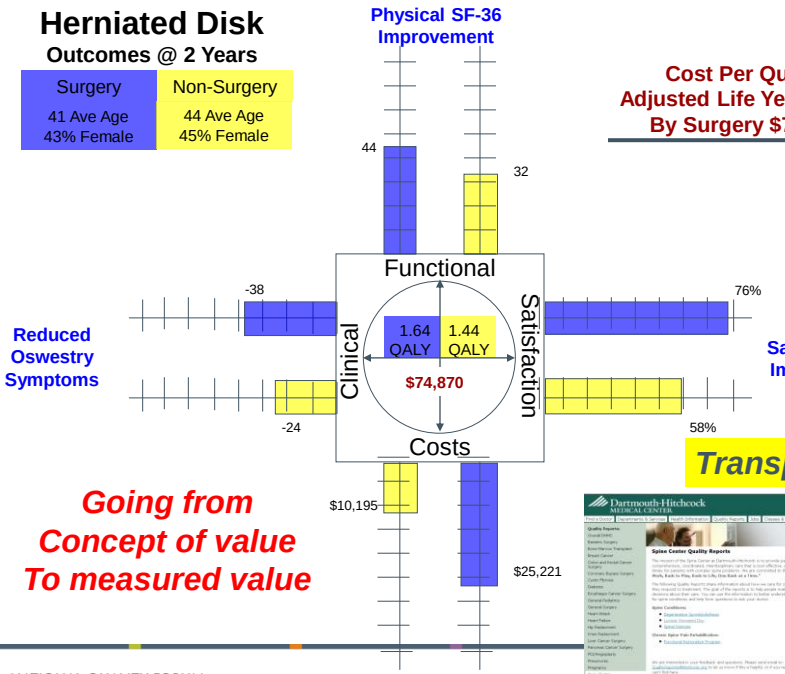
The summary report generated from patient-reported data is critical to a physician's ability to care for a patient



Herniated Disk

Outcomes @ 2 Years

Surgery	Non-Surgery
41 Ave Age 43% Female	44 Ave Age 45% Female

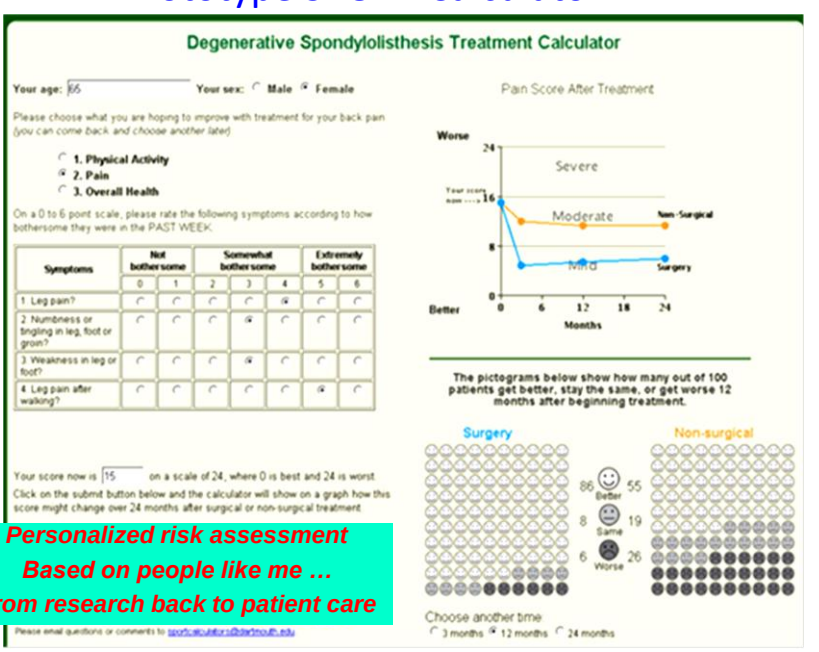


Cost Per Quality Adjusted Life Year Added By Surgery $74,870

Transparency

Going from Concept of value To measured value

Prototype SPORT Calculator



Personalized risk assessment Based on people like me … From research back to patient care

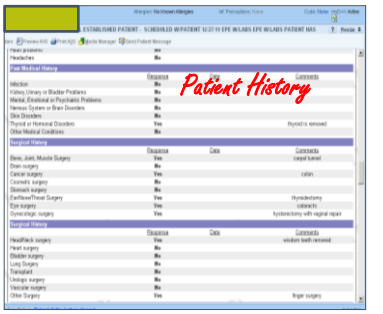
PRIM in eDH

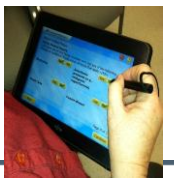
Patient Reported Information & Measures

Advantage of Dartmouth-Hitchcock's model of integrating patient-reported data into care

Patient Care	• Patient and provider engagement • Whole patient care • Informed patient choice
Research	• Research as part of clinical practice • Same system for practice and research • Comparative effectiveness research • Patient-centered, value-based research
Health System	• Patient-reported outcomes reporting • More efficient, complete visit documentation • Practice improvement based on outcomes • Value-based payment measures for ACOs*

**Value-based payment measures will be used for Accountable Care Organizations (ACOs), future reimbursements around episode bundled measures*





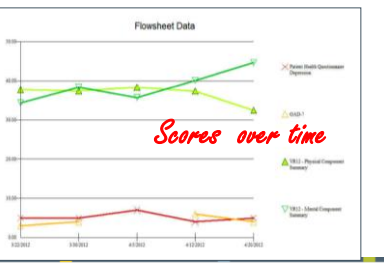
SUCCESSES

Early Adopters

- Primary Care - GIM*
- Pain Center
- Orthopedics (Wrist, Hip & Knee*, Shoulder)
- Plastics
- Urology
- Vascular
- Compre
- Neuro (Oncology, Multiple Sclerosis)
- Spine Center
- Occupational Medicine

**Integrates*

18 patient populations



Summary: to improve outcomes & value must capture PROs in flow of care and use it to improve outcomes and value of care for individual patients and populations

Lessons

- Patients' reaction: 84% positive*
  - "Visit became very helpful, thorough & informative"
- Providers' reactions*
  - "Patients get more involved in their care."
  - "This changes how care is delivered."
- Sustainable & replicable
  - 10 years at Spine Center & 18 DH programs & 70,000 patients
  - 13 SPORT sites & > 20 other health systems

Recommendations

- Successful feed forward PROs use design principles
  - Fit PROs into care flow to make it easier for patients and providers to do right thing
  - Co-design with stakeholder input for best end-user utility
  - Educate patients and providers on how to use PROs: providers must pay attention to patient's data
  - Capture data from other sources to improve utility of information
  - Continuously improve PRO system based on user's experiences & new technology

*Cite: Hvitfeldt H, Carli C, Nelson EC, Mortenson DM, et. al. Feed Forward Systems. *Quality and Safety in Health Care*, 18(4); 247-256, October-December 2010.

Using Patient-Reported Information to Improve Health Outcomes and Health Care Value:

- CASE STUDIES FROM DARTMOUTH, KAROLINSKA AND GROUP HEALTH
- Click Here to Download

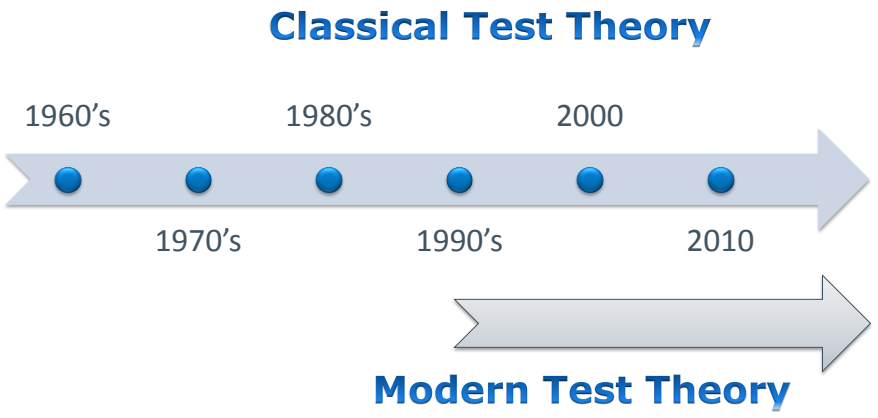
Methodological Issues: Method of Administration/Collection &Response

Lewis Kazis, Sc.D
Boston University School of Public Health

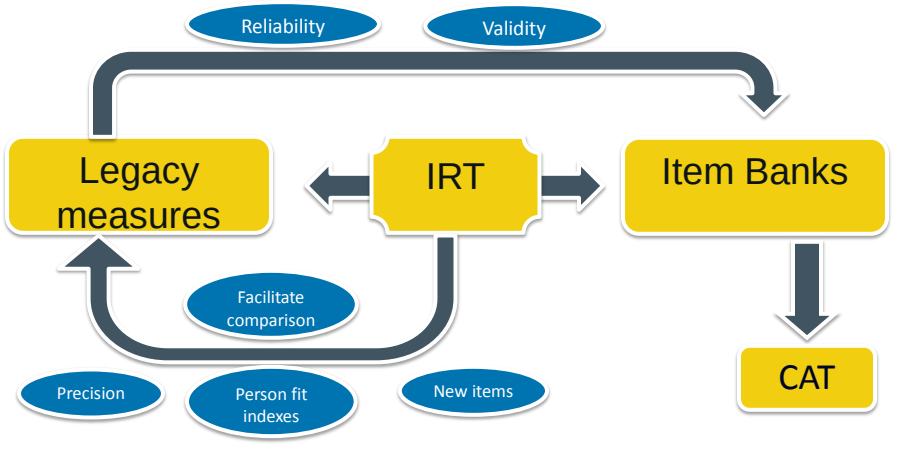
OVERVIEW

- **Historical Development of Short Form Assessments**
- **Bridging measurement tools:**
  - Legacy/Static measures
  - IRT/CAT measures
- **Mixed Mode Approaches**
- **Missingness / Bias**

Timeline
Classical and Modern Test Theory



Binding Framework



(1) Legacy / Static Measures

- Test score = true score + error score.
- The observed score is item sample dependent, and the statistics are respondent sample dependent.
- Longer tests are more reliable than shorter tests.

(2) Legacy / Static Measures

- Meaningful scale scores are obtained by comparisons of position in a score distribution.
- Interval scale properties are achieved by selecting items that yield normal raw score distributions.

(1) Item Response Theory (IRT) / (CAT)

- A set of generalized linear models that connect observed survey responses to a subject's on an unmeasured underlying latent construct.
- Assume unidimensionality (the scale items solely measure one construct).
- Assume uncorrelated items on a scale.
- Shorter tests can be more reliable than longer tests.

Legacy / Static Measures

Advantages	Disadvantages
Extensively tested for reliability and validity across multiple settings and populations. Fewer resources needed to implement, compared to CAT. The expertise to implement them is matured. Can be integrated with new technology (internet)	The time to complete the instrument is usually longer than CAT. Instruments are less flexible to update and recalibrate, compared to CAT. Requires larger samples to avoid spurious results.

(1) Item Response Theory (IRT) / (CAT)

Advantages	Disadvantages
Estimate person level traits within subset of items. Usually requires smaller sample sizes. Less vulnerable to floor an ceiling effects.	DIF calculation maybe problematic for multidimensionality assessment (prob. of responding in different cat. vary across different subgroups given equiv. levels of underlying attribute). Requires front end technology to implement the instruments. Additional assistance is usually necessary to facilitate successful patient-technology interaction.

(2) Item Response Theory (IRT) / (CAT)

Advantages	Disadvantages
Useful in assessing change. Greater precision of measurement. .	High startup costs. Software and hardware is commonly proprietary and expensive. .

Improving Legacy measures using IRT methods

- In a number of cases, legacy measures represent the foundation for CAT and item banks development.
- Item banks calibration adequately identify problematic legacy wording, enabling the enhancement of legacy measures in terms of reliability and validity.

References.
Hays RD, Morales LS, Reise SP. Item Response Theory and Health Outcomes Measurement in the 21st Century. *Medical care*. 2000;38(9 Suppl):II28.
Fries JF, Krishnan E, Rose M, Lingala B, Bruce B. Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory. *Arthritis Res. Ther.* 2011;13(5):R147.

Improving Legacy measures using IRT methods

Facilitate the development of new items to improve existing measures

- Legacy HAQ, SF-36, PF-10, have been improved using PROMIS:
  - Present tense
  - Five-item response categories
  - Improved quality and phrasing

References.
Hays RD, Morales LS, Reise SP. Item Response Theory and Health Outcomes Measurement in the 21st Century. *Medical care*. 2000;38(9 Suppl):II28.
Fries JF, Krishnan E, Rose M, Lingala B, Bruce B. Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory. *Arthritis Res. Ther.* 2011;13(5):R147.

Improving Legacy measures using IRT methods

Facilitate comparison across indexes

- Fryback and colleagues found among 5 utility scales, that each measurement identified health in a very similar fashion and are approximately linearly related.
- However, death remains controversial, and its location varied across scales.

Reference.
Fryback DG, Palta M, Cherepanov D, Bolt D, Kim J-S. Comparison of 5 health-related quality-of-life indexes using item response theory analysis. *Med Decis Making*. 2010;30(1):5–15.

Modes of Administration

- **Face to face interaction.**
- **Self administration (paper and pencil).**
- **Telephone.**
- **Computer-based assessment.**

Modes of Administration

- **Face to face interaction.**
  - Overall, responses give a more optimistic picture of health, compared to self-administration
  - White coat effect? It may be related with the positive effect of human involvement (rapport)
  - Interviewees overwhelmingly preferred it over the other modes.

Modes of Administration

- **Self administration**
  - Self-report is accurate (fewer sources are variation).
  - Lower scores (worse health) are usually reported, compared to face to face modes.
  - Less expensive than face to face interviews
  - Anonymity may yield more accurate rates for the “socially undesirable behavior”

Modes of Administration

- **Telephone**
  - Lower response rate compared to personal interviews. However, it costs less than half of the latter.
  - Less sensitive to non-response bias, compared to mail surveys.
  - Problematic for older adults (higher prevalence of hearing impairments).
  - Preferred over self administered surveys for individuals with lower literacy levels.

Modes of Administration

- **Computer-based assessment.**
  - Tailored “real time” results, immediately available to users and providers.
  - High rates of acceptance, even among interviewees without previous experience with computers.
  - The missingness of data may be reduced.
  - May capture data more accurately for “socially undesirable behaviors”.

Modes of Administration

- **Electronic vs. Paper instruments.**
  - Both instruments are comparable.
  - A critical review of 56 studies found average correlations exceeding 0.90 between electronic and paper assessments.

Reference: Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health*. 2008;11(2):322–333.

Missingness / Bias

- **Lower response rates**
  - Response rates are systematically declining over time.
  - Not enough evidence to determine the potential effects of unit non-response.
- **Response bias**
  - Greatly limits the generalizability of survey findings.
  - Homogeneous populations are less affected by response bias

Missingness / Bias

- **Imputation of missing values**
  - Missing items
    - Simple mean imputation (should satisfy many conditions first)
    - General imputation methods (e.g. GEE).
  - Missing forms
    - multivariate repeated measurements (analysis of variance)
    - modified regression estimates, (MRI estimator).
    - Random effects.

Missingness / Bias

- **Discerning Unbiased survey findings**
  - Estimating response rates
    - 60% as an acceptability “rule of thumb”.
  - Evaluating non response bias
    - More difficult to assess than response bias.
    - However, the representativeness of the sample should be assessed somehow.
  - Reporting non response rate.

Summary

- Hybrid approaches necessary that bridge Legacy and IRT/CAT approaches for purposes of application to systems for measurement performance
- Mixed mode approaches are necessary so that flexibility in the protocols is possible in real world settings.
- Missing data is a fact of life in real world settings and adjustments for missingness is required to adjust for bias in results.

Methodological Issues: Method of Administration/Collection &Response

NATIONAL QUALITY FORUM

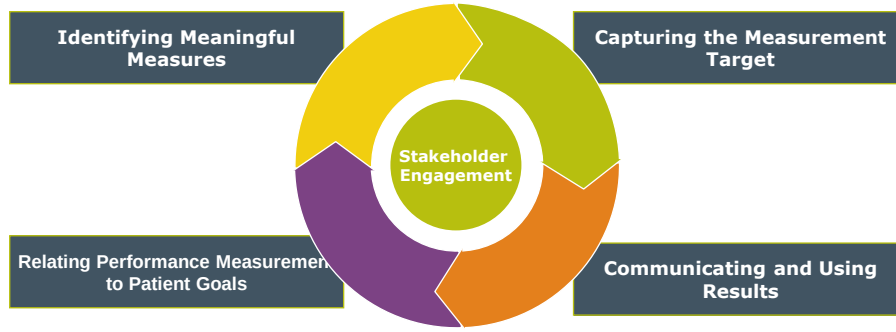
Lori Frank, PhD
Patient Centered Outcomes Research Institute, Washington, DC

Why Engage?

Guiding Principles for Stakeholder Engagement

Why Engage?

Guiding Principles for Stakeholder Engagement in Performance Measurement



Methodological Issues: Selecting Patient-level PROs

NATIONAL QUALITY FORUM

Eugene C. Nelson, DSc, MPH
The Dartmouth Institute
Dartmouth-Hitchcock Health

When can general health status measures be utilized & when should condition-specific measures be utilized? Are there any setting specific issues for selection of PROs?

- In general, prudent to use both general and condition-specific measures of functional status and symptoms
  - Examples: Spine, heart failure, total joint replacement of hip and knee, depression
- Imperative to use general health status measures under some conditions
  - When patient has multiple comorbid conditions, e.g., 76.1% of heart failure patients have 2 or more chronic conditions
  - When screening for problems that may be important but can easily go undetected
    - CABG or Spine or AMI or post-partum: screen for mental health problems
    - Annual Wellness Visit or periodic health exam: screen for functional problems and high health risk status
- Settings for PROs use: home, outpatient, inpatient, ECF (subject to patient's ability to provide data)

What conditions would be most sensitive to measuring changes in patient health status/outcomes? What is the variation in patient-level scores related to clinical interventions (e.g., hip replacement)?

- Function: Chronic conditions with large impact on physical, mental and role function such as heart failure, depression, ischemic heart disease, Parkinson's Disease, low back pain, osteoarthritis, rheumatoid arthritis, etc.
- Function: Surgical conditions with large impact on physical, mental and role function such as CABG, TJA, bariatric surgery, spine surgery, etc.
- Risk: People at high risk of avoidable death … health risk status measurement and monitoring using Framingham Index or all cause mortality index such as people with cardio-metabolic syndrome, hypertension, diabetes, hyperlipidemia, and high risk health-related behaviors, etc.

Table 1. Variation in PROs associated with selected interventions

Population	Measure/range	Average Change
Herniated disk: SPORT	SF-36 PCS / 0-100	44 surg v 32 non-surg
Stenosis: SPORT	SF-36 PCS / 0-100	17 surg v 17 non-surg
Spondylolythesis: SPORT	SF-36 PCS / 0-100	27 surg v 8 non-surg
Depression: EBM Protocol	PHQ-9 / 0-27	10 or greater = clinical depression v < 5 = remission or 5-9 = response
Carpal Tunnel: Trumble Trial	BCTQ/ 1-5	3.1 pre-surg v 1.8 post-surg
TJA-Hip: UK Knee Society	WOMAC/ 0-100	42 pre-surg v 70 post-surg
RA: Sweden Registry	DAS / 0-10	5.0 to 2.8 at 1 yr
Aortic Valve Stenosis: PARTNER Trial	KCCQ/0-100	32 in Transcatheter Aortic Valve Replacement (TAVR) vs 4 in Meds at 1yr
Angina: COURAGE Trial	SAQ Angina Frequency (AF) & QoL/0-100	AF: 68 vs. 87 baseline v 1 yr post Percutaneous coronary intervention (PCI) QoL: 51 vs. 76 baseline v 1 yr post PCI --- AF: 87 in PCI vs. 84 w/meds QoL: 76 in PCI vs. 73 w/meds

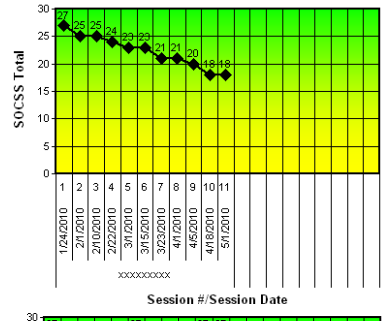
Appendix: References

- LefevreT,Kappetein AP,Wolner E,Nataf P,Thomas M,Schachinger V,De Bruyne B, Eltchaninoff H, Thielmann M, Himbert D, Romano M, Serruys P, Wimmer-Greinecker G; on behalf of the PARTNER EU Investigator Group. One year follow-up of the multi-centre European PARTNER transcatheter heart valve study. *Eur Heart J.* 2010;32:148–157.
- MacDonald SJ, Charron KD, Bourne RB, Naudie DD, McCalden RW, Rorabeck CH. The John Insall Award: Gender-specific total knee replacement: prospectively collected clinical outcomes. Clin Orthop Relat Res. 2008;466. Doi:10.1007/s11999-008-0430-1.
- Nelson EC, Hvitfeldt H, Reid R, et al. (2012) Using Patient-reported Information to Improve Health Outcomes and Health Care Value: Case Studies from Dartmouth, Karolinska and Group Health. Lebanon, NH: The Dartmouth Institute for Health Policy and Clinical Practice.
- Spertus J, Peterson E, Conard MW, et al; Cardiovascular Outcomes Research Consortium. Monitoring clinical changes in patients with heart failure: a comparison of methods. *Am Heart J.* 2005;150(4): 707-715.
- Trumble TE, Diao E, Abrams RA, Gilbert-Anderson MM. Single-portal endoscopic carpal tunnel release compared with open release: a prospective, randomized trial. J Bone Joint Surg Am 2002;84:1107-15.
- Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. JAMA 2006;296:2441-50.
- Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, Herkowitz H, Cammisa F, Albert T, Boden SD, Hilibrand A, Goldberg H, Berven S, An H; SPORT Investigators. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 2008; 358:794-810.
- Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 2007;356:2257-70.
- Weintraub WS,Spertus JA,Kolm P,Maron DJ,Zhang Z,Jurkovitz C,Zhang W, Hartigan PM, Lewis C, Veledar E, Bowen J, Dunbar SB, Deaton C, Kaufman S, O'Rourke RA, Goeree R, Barnett PG, Teo KK, Boden WE; for the COURAGE Trial Research Group. Effect of PCI on quality of life in patients with stable coronary disease. *N Engl J Med.* 2008;359:677–687.

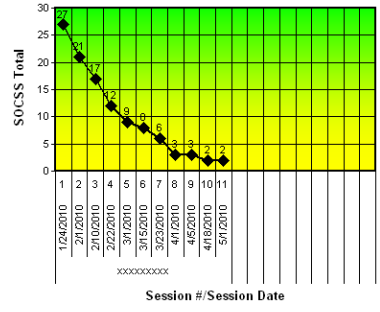
Key Considerations for Incorporating PROs into Electronic Health Records

Uma Kotagal, MSc
Cininnati Children's Hospital Medical Center

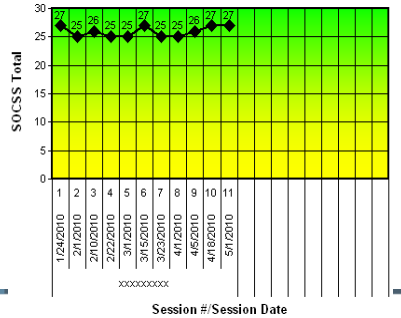
NATIONAL QUALITY FORUM



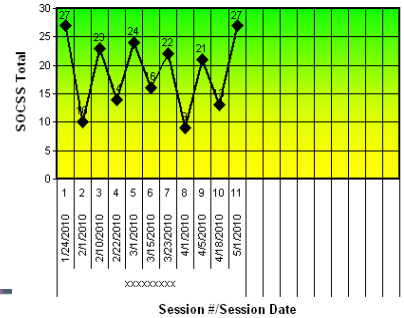
Slow Decline



Rapid Decline

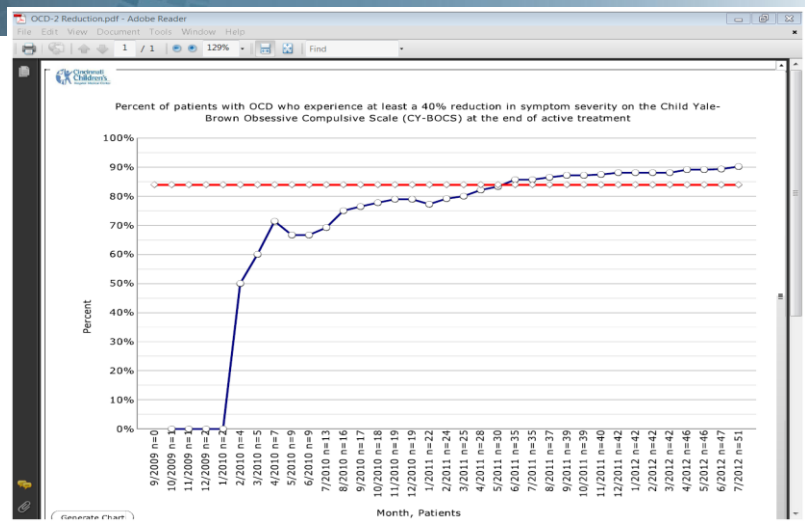


No Improvement



Saw Tooth

Overall reduction in symptoms by 40% in most patients as seen on CYBOCS



Recap of Day 1
Overarching Themes

(working draft)

NATIONAL QUALITY FORUM

Overarching Theme: Person-Centered

- Terminology: “person” versus “patient”
- Patient experience not just with the healthcare delivery system; includes whether needs are met and linked to other services to improve health
- Patient important outcomes: relevant and meaningful to persons and their families (vs. research)
- PROs must be actionable to persons, providers, policy makers, others. Add to list of essential characteristics
- PROs are an important step towards engaging patients and providers in creating a person-centered environment

Overarching Theme: Accountability

- PROs and state of readiness for purposes of accountability
  - What is the pathway to accountability measures?
  - As measures expand beyond sickness/illness to health/well-being shared accountability will be required beyond the healthcare system

Lessons from the Field

- Person buy-in:
  - Persons feeling "spammed" by survey requests
  - Engage persons in determining what PROs are meaningful to them
- Health Professional buy-in
  - Fitting results into the workflow
  - Knowing what to do with the results

Lessons from the Field

- Guiding principles for stakeholder engagement in performance measurement – engage patients at all steps
  - Identify meaningful measures
  - Capture the measurement target
  - Communicating and using results
  - Relating performance measures to patient goals

Approaches to Implementation

- Just get started! Let providers innovate.
- Approaches put forth:
  - Initially measure that PROs are collected (e.g., process measure) on relevant patients recognizing outcomes are more meaningful [getting people used to it] NEEDS SHARPENING
  - Usability, feasibility, actionability are paramount to selection of PROs –determine first before implementing into accountability programs
  - Start with focused areas where we have validated measures (e.g., hip heart) and have good evidence on how to improve
  - Generic assessments offering the "biggest bang for the buck"

Themes: Key Methodological Issues

- Missing data
  - Safeguarding against excluding sicker patients
  - Bias introduced by how the tool is administered
  - Engagement strategy needed over time (e.g., response rate)
- Need bridges that combine use of legacy tools and Item Response Theory -- advantages/disadvantages to both suggest hybrid aprroach

Themes: Key Methodological Issues

- Reconciliation is needed around heterogeneity of multiple approaches (use of different tools & modes of administration) for comparability to make sure they are equivalent
- “Leading” measure versus “lagging” measures (e.g., mortality doesn’t have a guideline)
- Outcomes with high face validity are not required to be based on guidelines (NQF Evidence Task Force)

Themes: Electronic Health Records

- PROs and parsimony
  - Building blocks that can be leveraged for different purposes (e.g., “app” store)
  - Flexible platform (e.g., PROMIS)
  - Infrastructure that exists which can accommodate new tools (e.g. reusable codes)
  - Accommodating multiple styles respectful of ways patients wish to engage with the system

Electronic Health Records

- New reality: Patient will “own” the record and provide and extract information –implications for existing EhRs
- Patient experience is still needed but there are risks putting into EhR (e.g., recrimination)
- IP & copyright issues for codes & instruments (not limited to EhR)

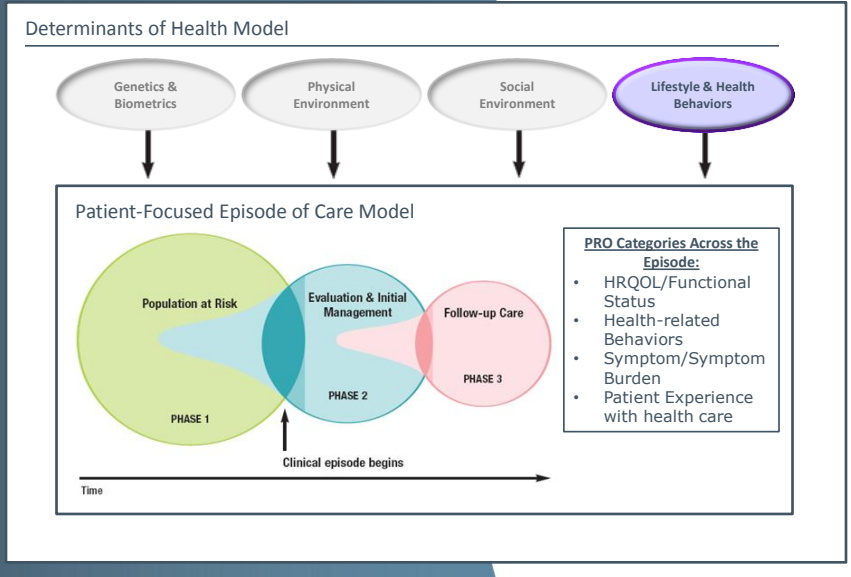
Patient-Reported Outcomes Workshop #1

Breakout Session

July 31, 2012

NATIONAL QUALITY FORUM

Framing PROs Within Existing Conceptual Models



AGENDA

- 9:35-9:45 Round Robin Introductions & identify spokesperson
- 9:45-10:30 Discussion of Question #1
  - What characteristics should be used to identify PROs for potential use in performance measures? Will these differ based on the needs of the end-user?
- 10:30-11:15 Discussion of Question #2
  - What existing individual-level PROs have these identified characteristics and are candidates for potential development of performance measures?
- 11:15-11:30 Synthesis and complete templates for report out

Patient-Reported Outcomes Workshop #1

Breakout Session Report Back

July 31, 2012

(working draft)

NATIONAL QUALITY FORUM

Report Out: HRQoL/Functional Status
Question #1: Characteristics to Select PROs

- HRQoL/Functional Status
- **Adequate measure properties (scientifically and clinically defensible)**
  - Conceptual and Measurement Model
  - Reliability
  - Validity (and consideration of validity for proxy response; link to alternative modes)
    - Face validity (clinical/patient sensibility) w/ respect to how relevant it is to patients & clinicians
    - Risk-adjustable
  - Responsiveness
  - Burden
    - Literacy level and cognitive demand
    - Something that be practically incorporated into clinical practices
- **Meaningfulness to clinicians and other health professionals, patients, and systems**
  - Evidence for usefulness / appropriateness for specific pops
  - Interpretability of Scores
- **"Implementabilty"**
  - "Electronicability"/alternate modes
  - Translatable
  - "Game-ability"
  - Unintended consequences
  - "Proprietariness"

Report Out: HRQoL/Functional Status
Question #2: PROs with the Identified Characteristics

- HRQoL/Functional Status
- Top 3-5 candidates for potential development of performance measures (attached slide color coded **green** for in use)

1. Generic
2. Disease specific

HRQoL/Functional Status
Question #2 Notes for reference
PROs with the Identified Characteristics – Generic

- **AMPAC**
  - Mobility and Self-Care
- **VR12 (med adv)/VR36/VR6D**
- BRFSS (three QoL – healthy day items)
- **PROMIS** (**adult** and **children**)
- **EQ5D and HUI** (utility); proprietary
- Sickness Impact Profile
- Money Follows the Person QoL Scale
- **Basic and instrumental scale of daily living**
- **Social and productive activity scales**
- QWB
- Child Health questionnaire
- **CHIP**
- **Kids Screen**
- **FIM (follow-up)**
- **Community Integration questionnaire**
- SEIQoL
- **EVGFP**
- **WHO QoL**
- Penny E
- WHO DAS (DALYS)
- **SF family**
- **Restricted Activity Days**

HRQoL/Functional Status
Question #2 Notes for reference
PROs with the Identified Characteristics

- **PHQ-9**
- **CESD**
- **VF-14** (visual functioning)
- **SGRQ**
- **CRQ**
- **KDQoL**
- **Oxford Knee Score**
- **ODI**
- **HAM-D**
- **MADRS**
- **FACT/FACIT**
- **EORTC QLQ**
- **FLIC**
- **Kansas City Heart Failure**
- **Minnesota living with heart failure**
- FAHI
- **Arthritis questionnaire (HAQ)**
- **ACT**
- **MOS-HIV**
- **IPSS**
- AIMS
- **BPI (pain)**
- **IIEF**
- **CHART**
- BOQ
- **PRO CTCAE**

Report Out: Health-Related Behaviors
Question #1: Characteristics to Select PROs

- Health-Related Behaviors
- Synthesis of top 3-5 characteristics to identify PROs for potential use in performance measures

1. *An **evidence-based** justification for selection suggesting a measure is **actionable** to appropriate end-users.
2. Degree of importance to adequately capture the **impact of a health-related behavior on a patient**
3. Assessment of the **level of accountability**; individual, culture, environment, resource accessibility, etc
4. Lends itself to a model of **shared decision making**; **engaging patients** in their own self-management & goal attainment

Report Out: Health-Related Behaviors
Question #2:PROs with the Identified Characteristics

- Health-Related Behaviors
- Top 3-5 candidates for potential development of performance measures

1. **Federally Sponsored Health Surveys**
   -BRFSS, NHANES, HOS, ACO CAHPS, Physical Activity FS
2. **Commercial Health Risk Appraisals**
   - Stay Well, Health Media, U of M,
3. **Behavioral Health & Substance Use**
   -PHQ-2, CAGE, Audit-C,
4. **Specific High Impact Health Related Behaviors**
   - How's Your Health, PROMIS, Smoking Index, Framingham Index

* Inclusion of pediatric category TBD

Report Out: Symptoms and Symptom Burden
Question #1: Characteristics to Select PROs

Symptoms and Symptom Burden
Synthesis of top 3-5 characteristics to identify PROs for potential use in performance measures

- Patient engagement
  - Identify important outcomes to patients
  - Involve in development, testing, use
  - Assure cultural, linguistic, literacy adaptability
- Purpose/Goal
  - Identify end users/stakeholders (patient, caregiver, provider, plan, payor)
  - Specify context of use (disease, population, time horizon, setting, interpretation of results)
  - Articulate conceptual/measurement model
  - Actionability
- Measurement properties
  - Content validity
  - Quantitative measurement properties (reliability, construct validity, sensitivity, appropriate recall period)
- Feasibility
  - Consider mode of administration, interoperability with HER
  - Burden to patients, providers, infrastructure requirements

Report Out: Symptoms and Symptom Burden Question #2: PROs to Consider for Measurement

Symptoms/Problems and Symptom Burden

- Pain
  - Worst pain item from the Brief Pain Inventory
- Interference with ADLs/IADLs
- Dyspnea
- Fatigue
- Mood
- Memory
- Sleep disturbance
- Cognitive disturbance
- Mobility
- Nausea/Vomiting
- Constipation
- Diarrhea
- Continence
- Sexual dysfunction
- Appetite loss/anorexia
- Edema
- Body image
- Sensory loss

Report Out: Patient Experience with Healthcare Question #1: Characteristics to Select PROs

- Goal: Moving beyond past episode of care: Knowing what happens next, and what patient's role is - actionable and responsive (i.e., questions that ask about state of patient and ensuing action)
- Principles:
  - Continuum of care or longitudinal
  - Not just provider, setting or episode/encounter-specific
  - **Person-centered**

Report Out: Patient Experience with Healthcare Question #1 Characteristics to Select PROs

- Needs to be actionable that leads to change/improvement by unit of analysis: provider/system as well as the person
- Should be linked to the individual's goals, which encourages engagement but also flexible enough to account for population
  - Example: PROMIS data bank where items can be selected based on person needs/goals and provider and system needs (setting, etc.)
- Minimize provider/system and patient burden
- Needs to accommodate cultural and language preferences
- Needs to be responsive to individual preferences
  - Include alternative methods of administration (including who, where and when)

Report Out: Patient Experience with Healthcare Question #2 : Existing Tools

- CAHPS (acute/amb care) (NQF-endorsed)
  - provider communication; access to care; timeliness to care; care coordination; patient and family involvement; language access; shared decision-making; care transition; cultural competency; staff helpfulness; experience of environment; alternative medicine; communication about medicine (overlap w/health literacy); pain management
- VA FATE survey (NQF-endorsed)
  - captures episode – family support; care assessment (degree)

Report Out: Patient Experience with Healthcare Question #2 Existing Tools

- National Core Indicators (long-term care)
  - in use by 35 states; nationwide in 3 years
  - **identifies individual responses across multiple areas: family outcomes; individual survey and 3 family surveys; family involvement; health and welfare; therapeutic interventions; medications and incidence; systems issues; safety (home); service coordination; staff stability**
  - over 100 individual performance measures (some risk adjusted)
  - 2 sections (consumer section)
  - Who and how they are administering survey also important.
  - How do you balance burden concerns?
    - Background done by case manager; staff input and also receive consumer perspective. Mission of community-based org, consumer engagement throughout

Report Out: Patient Experience with Healthcare Question #2 Existing Tools

- Patient Activation Tool – shared decision-making; power dynamic; how person felt from encounter; used in chronic care management
  - Concern - measuring output or input?
  - "Self confidence" (suggested domain to measure) – outcome of positive experience. Immediate metric between encounter and outcome
  - Hibbard and Colleagues may have a tool to measure self confidence

Report Out: Patient Experience with Healthcare Question #2 Suggested, New or in Development

- Pacific Business Group on Health
  - Testing new tool on self care management; health status impact; shared decision-making
  - How much consumers pay for services – how they rate their health plan and not provider
- Clinician advocacy on behalf of new patients
- Suggested process measure – did system ask about goals for the visit? Did they respond/fulfill goals?
- Current - Cash and counseling (demo program from CMS) – persons with severe physical disabilities given budget and purchase own services; model worked in 3 states – person make own decisions; current assessment is qualitative, not standardized
- Self directive services – all states offer these options