# Measure Applications Partnership

## Dual Eligible Beneficiaries Workgroup Web Meeting

NATIONAL QUALITY FORUM

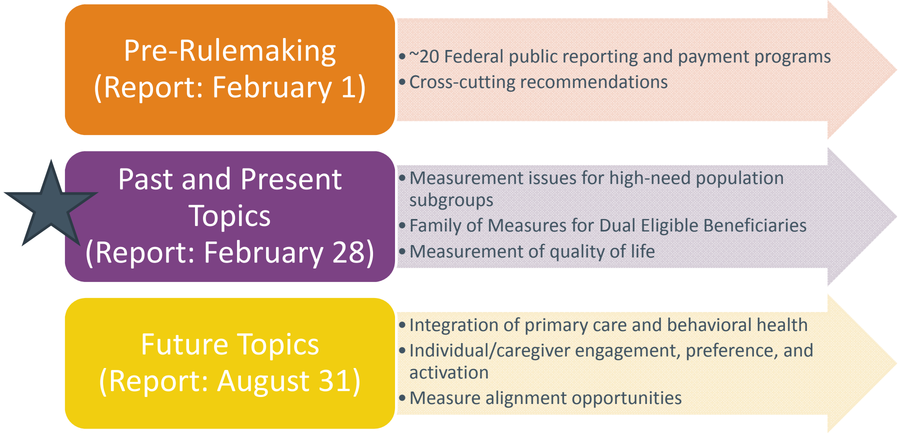
*November 21, 2013*

## *Welcome*

## The Project Year Ahead: Themes

- NQF and MAP will seek to understand and help improve:
  - Current issues in healthcare quality that are unique to dual eligible beneficiaries
  - Real-world issues encountered while implementing performance measures
  - Dialogue between stakeholders that use measures and stakeholders that develop and maintain them
  - Opportunities for improved measure alignment across programs and across care settings

## The Year Ahead: Content and Timing of Reports



**Pre-Rulemaking (Report: February 1)**
- ~20 Federal public reporting and payment programs
- Cross-cutting recommendations

**Past and Present Topics (Report: February 28)**
- Measurement issues for high-need population subgroups
- Family of Measures for Dual Eligible Beneficiaries
- Measurement of quality of life

**Future Topics (Report: August 31)**
- Integration of primary care and behavioral health
- Individual/caregiver engagement, preference, and activation
- Measure alignment opportunities

## Today's Meeting Objectives

- Explore similarities and differences between surveys and performance measures using the NQF framework for Patient-Reported Outcomes (PROs) in performance measurement
- Review selected resources currently available for assessing quality of life
- Engage workgroup members in a discussion of the challenges and opportunities related to measuring quality of life in the dual eligible beneficiary population

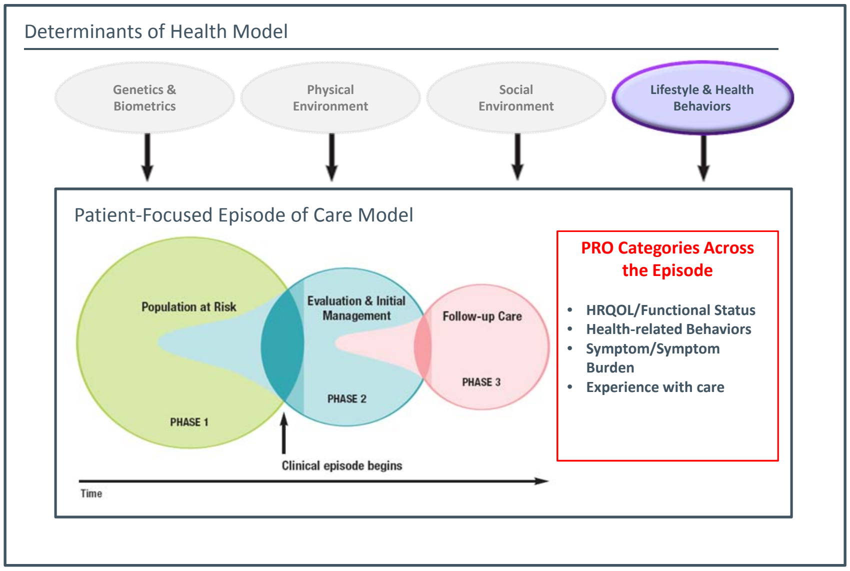
## *Perspective on Quality of Life Measurement from CMS*

# *NQF Framework on Patient-Reported Outcomes in Performance Measurement*

## Patient-reported Outcome (PRO)

- The concept of any report of the status of a patient’s health condition that comes directly from the patient ***without interpretation of the patient’s response by a clinician or anyone else***.
- PRO domains:
  - health-related quality of life including functional status
  - symptom and symptom burden
  - experience with care
  - health-related behaviors

## Framing PROs Within Existing Conceptual Models



Determinants of Health Model

Genetics & Biometrics

Physical Environment

Social Environment

Lifestyle & Health Behaviors

Patient-Focused Episode of Care Model

Population at Risk

PHASE 1

Evaluation & Initial Management

PHASE 2

Follow-up Care

PHASE 3

Clinical episode begins

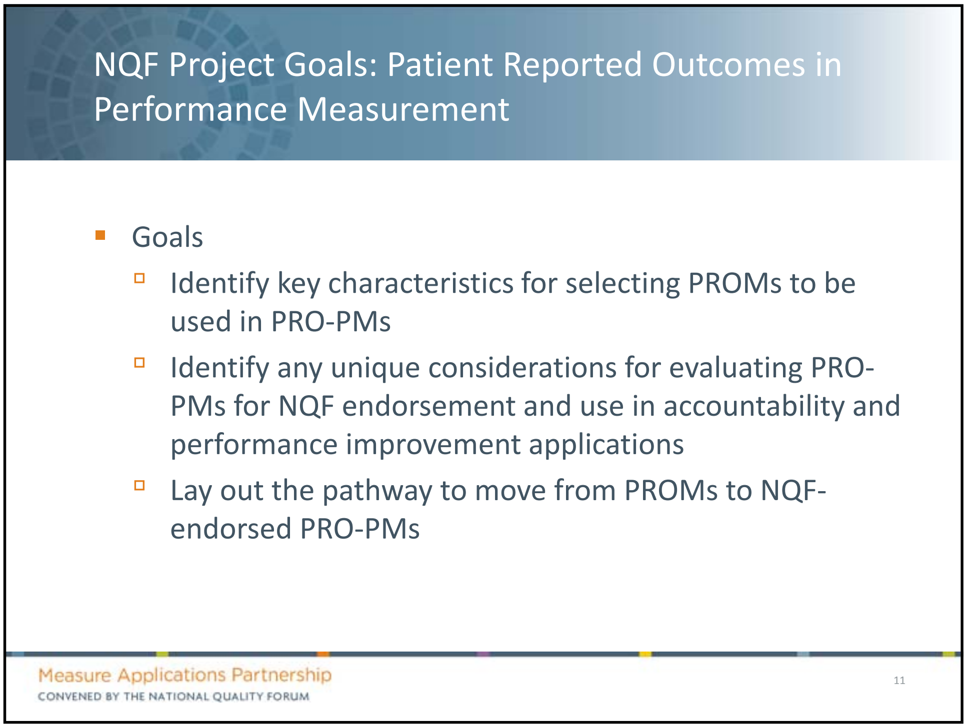
Time

**PRO Categories Across the Episode**

- HRQOL/Functional Status
- Health-related Behaviors
- Symptom/Symptom Burden
- Experience with care

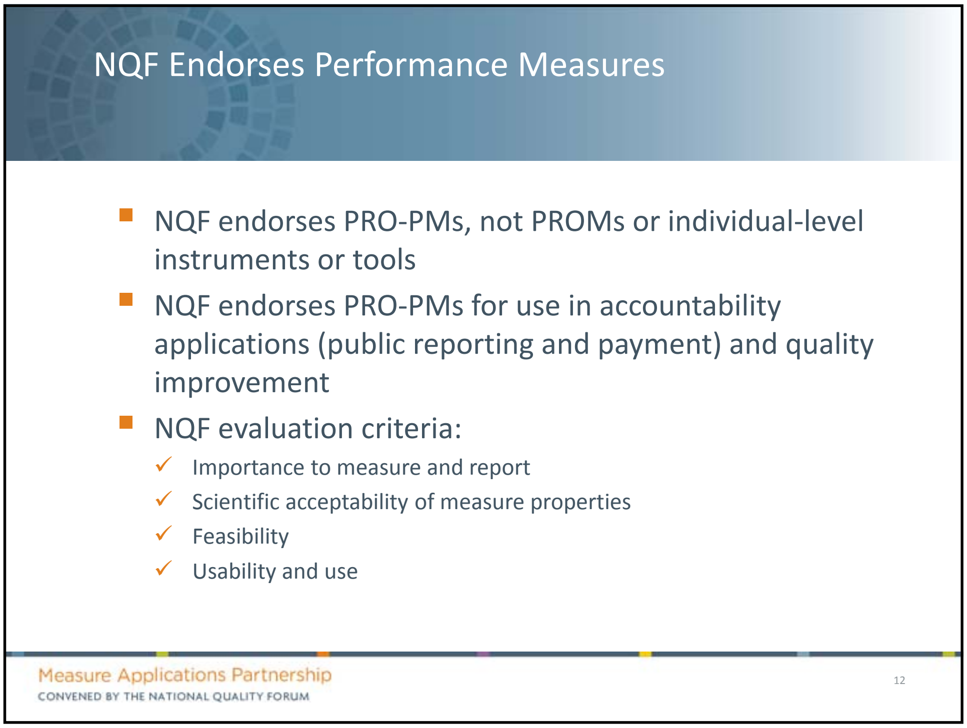
## Distinctions Among PRO, PROM & PRO-PM

| | Definition | Example: Patients With Clinical Depression |
|---|---|---|
| **PRO** **(patient-reported outcome)** | The concept of any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else. | Symptom: depression |
| **PROM** **(instrument, tool, single-item measure)** | Instrument, scale, or single-item measure used to assess the PRO concept as perceived by the patient, obtained by directly asking the patient to self-report (e.g., PHQ-9). | PHQ-9©, a standardized tool to assess depression |
| **PRO-PM** **(PRO-based performance measure)** | A performance measure that is based on PROM data aggregated for an accountable healthcare entity (e.g., percentage of patients in an accountable care organization whose depression score as measured by the PHQ-9 improved). | Percentage of patients with diagnosis of major depression or dysthymia and initial PHQ-9 score >9 with a follow-up PHQ-9 score <5 at 6 months (NQF #0711) |



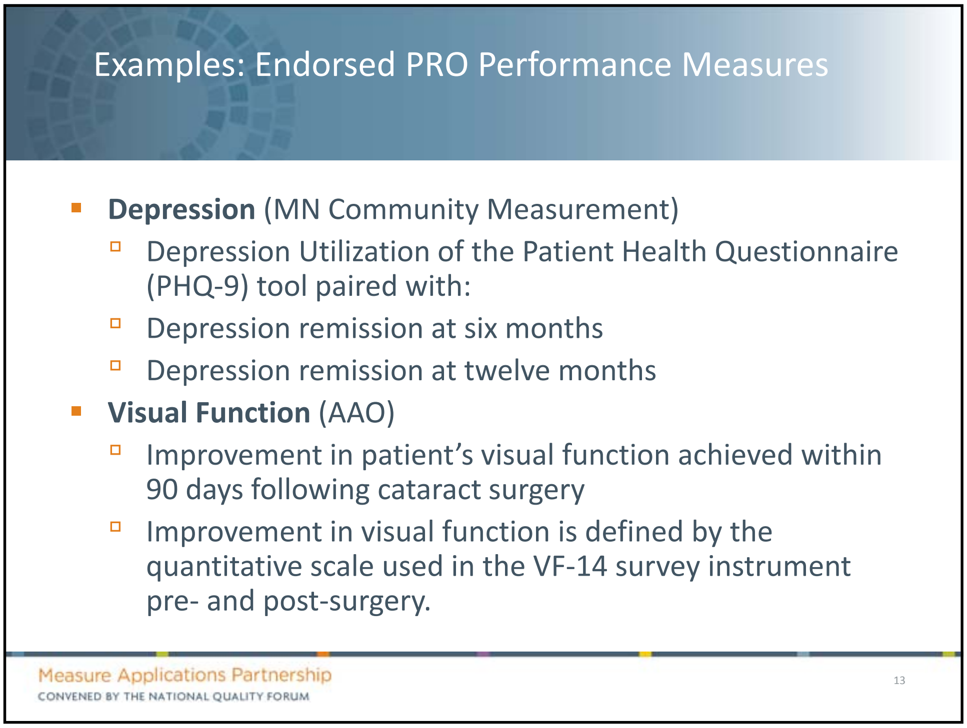
## NQF Project Goals: Patient Reported Outcomes in Performance Measurement

- Goals
  - Identify key characteristics for selecting PROMs to be used in PRO-PMs
  - Identify any unique considerations for evaluating PRO-PMs for NQF endorsement and use in accountability and performance improvement applications
  - Lay out the pathway to move from PROMs to NQF-endorsed PRO-PMs



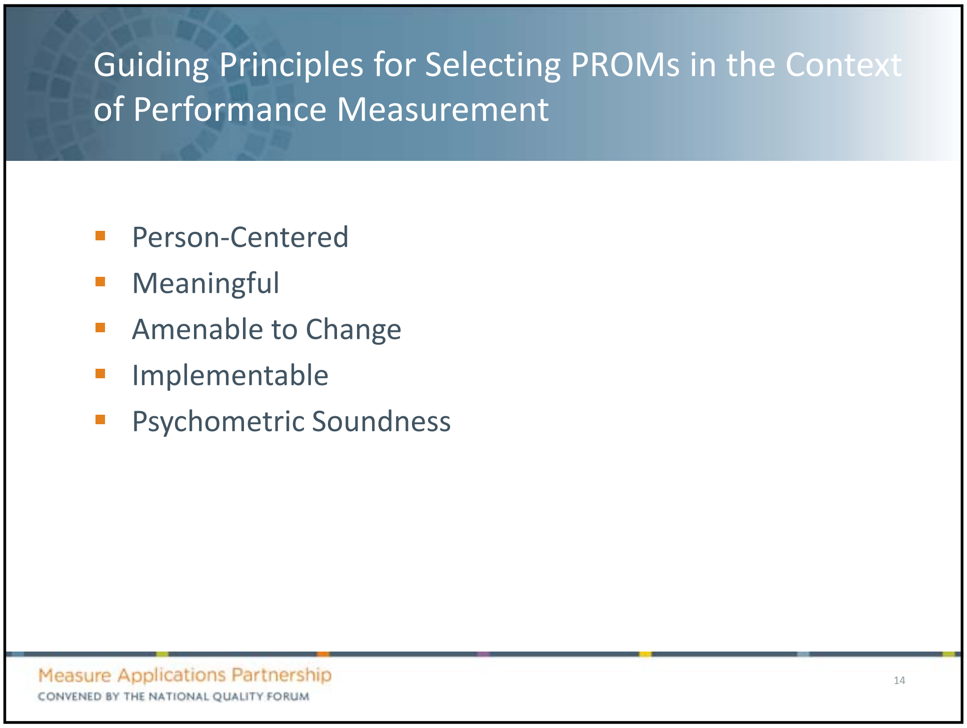
## NQF Endorses Performance Measures

- NQF endorses PRO-PMs, not PROMs or individual-level instruments or tools
- NQF endorses PRO-PMs for use in accountability applications (public reporting and payment) and quality improvement
- NQF evaluation criteria:
  - ✓ Importance to measure and report
  - ✓ Scientific acceptability of measure properties
  - ✓ Feasibility
  - ✓ Usability and use



## 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.



## Guiding Principles for Selecting PROMs in the Context of Performance Measurement

- Person-Centered
- Meaningful
- Amenable to Change
- Implementable
- Psychometric Soundness

## Person-centered

- Terminology: “person” versus “patient”
- PROs are an important step towards engaging patients and providers in creating a person-centered environment
- PROMs can be used to facilitate shared decision-making
- Assessment of patient experience should include whether needs are being met and linked to other services to improve health—not only healthcare delivery system focused

## Meaningful

- Patient ***important*** outcomes
- Relevance and degree of importance of the concepts measured by the PROM from the perspective of patients, their families and caregivers, as well as the clinicians and other health professionals that serve them

## Amenable to Change

- Evidence that the outcome of interest is responsive to a specific healthcare or support service intervention
- Key end users including patients, providers and systems should be motivated by the PROM to lead to improvement

- Counter argument: Some outcomes are worth measuring that may not be amendable to change by providers—but patients value and can be used to make informed decisions (e.g., time to recovery)

## Implementable

- Ease of fielding is an important consideration and its relationship to burden to both patients and providers
- Barriers to adoption:
  - *Data collection and reporting*
  - *Health literacy of patients*
  - *Cultural competency of clinicians and service providers*
  - *Costs associated with using PROMS especially if not available in the public domain*
  - *Adapting PROMs to computer-based platforms or other formats*

## Psychometric Soundness

- Conceptual and Measurement Model Documented
- Reliability
- Validity
- Interpretability of Scores
- Burden
- Alternate Modes & Methods of administration
- Cultural and language adaptations
- Electronic health record capability

## Pathway from PRO to NQF-endorsed PRO-PM PRO

**1. Identify the quality performance issue or problem**
- Include input from all stakeholders including consumers and patients

**2. Identify outcomes that are meaningful to the target population and are amenable to change**
- Ask persons who are receiving the care and services
- Identify evidence that the outcome responds to intervention

**3. Determine whether patient-/person-reported information (PRO) is the best way to assess the outcome of interest**
- If a PRO is appropriate, proceed to step 4

## Pathway from PRO to NQF-endorsed PRO-PM
## PROM

**4. Identify existing PROMs for measuring the outcome (PRO) in the target population of interest**
- Many PROMs (instrument/ scale/single-item) were developed and tested primarily for research

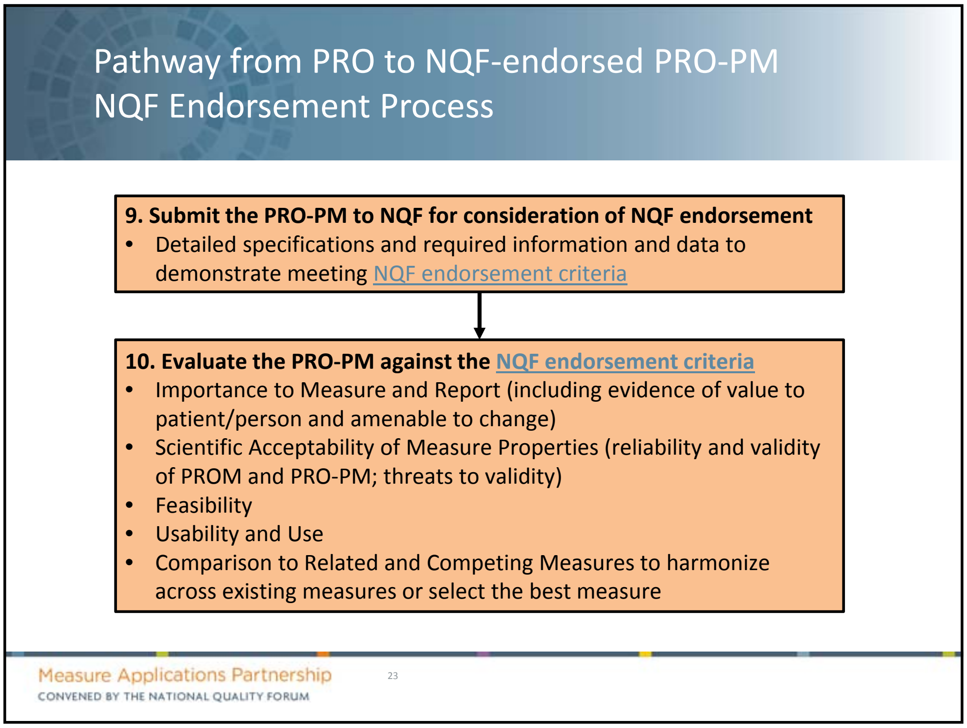
**5. Select a PROM suitable for use in performance measurement**
- Identify reliability, validity, responsiveness, feasibility in the target population

**6. Use the PROM in the real world with the intended target population and setting to:**
- Assess status or response to intervention, provide feedback for self-management, plan and manage care or services, share decision-making
- Test feasibility of use and collect PROM data to develop and test an outcome performance measure

## Pathway from PRO to NQF-endorsed PRO-PM
## PRO-PM

**7. Specify the outcome performance measure (PRO-PM)**
- Aggregate PROM data such as average change; percentage improved or meeting a benchmark

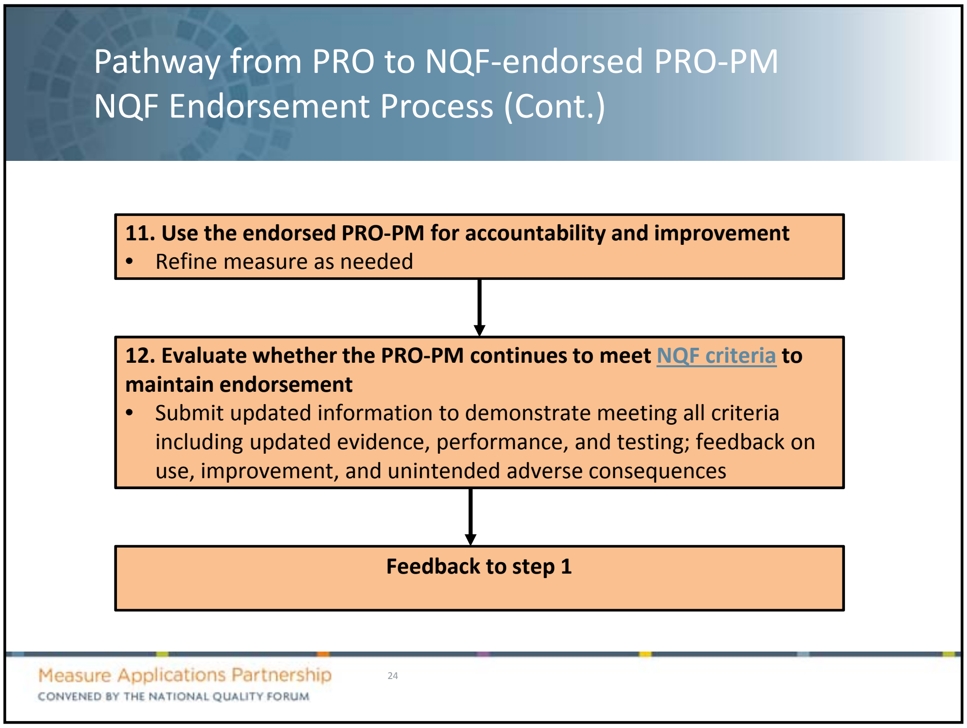
**8. Test the PRO-PM for reliability, validity, and threats to validity**
- Analysis of threats to validity, e.g., measure exclusions; missing data or poor response rate; case mix differences and risk adjustment; discrimination of performance; equivalence of results if multiple PROMs specified



## Pathway from PRO to NQF-endorsed PRO-PM
## NQF Endorsement Process

**9. Submit the PRO-PM to NQF for consideration of NQF endorsement**
- Detailed specifications and required information and data to demonstrate meeting NQF endorsement criteria

**10. Evaluate the PRO-PM against the NQF endorsement criteria**
- Importance to Measure and Report (including evidence of value to patient/person and amenable to change)
- Scientific Acceptability of Measure Properties (reliability and validity of PROM and PRO-PM; threats to validity)
- Feasibility
- Usability and Use
- Comparison to Related and Competing Measures to harmonize across existing measures or select the best measure



## Pathway from PRO to NQF-endorsed PRO-PM
## NQF Endorsement Process (Cont.)

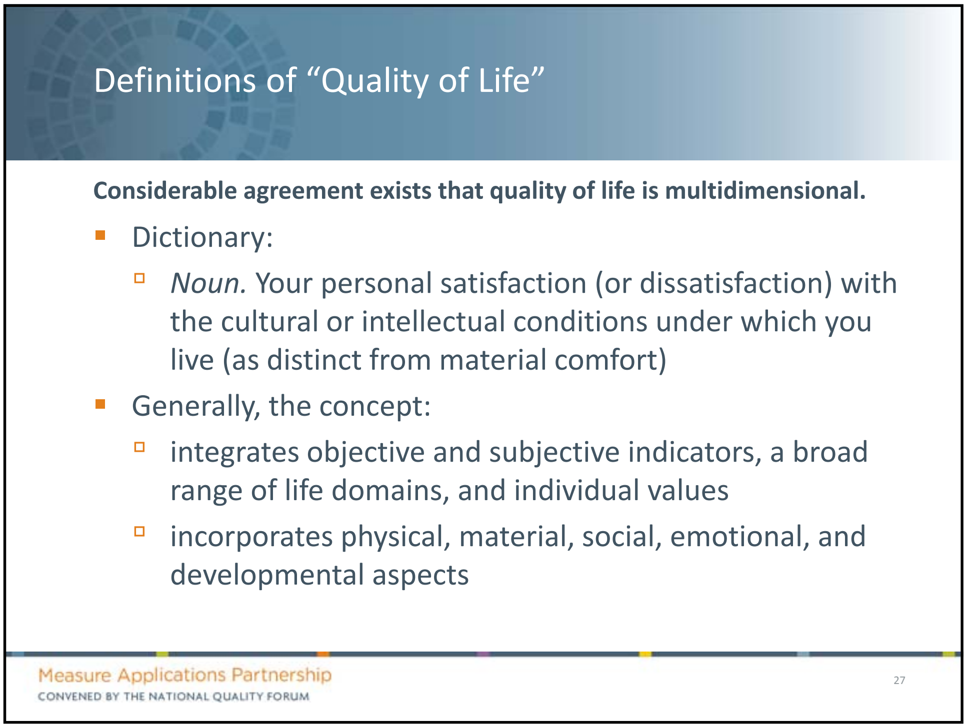
**11. Use the endorsed PRO-PM for accountability and improvement**
- Refine measure as needed

**12. Evaluate whether the PRO-PM continues to meet NQF criteria to maintain endorsement**
- Submit updated information to demonstrate meeting all criteria including updated evidence, performance, and testing; feedback on use, improvement, and unintended adverse consequences

**Feedback to step 1**

# Questions?

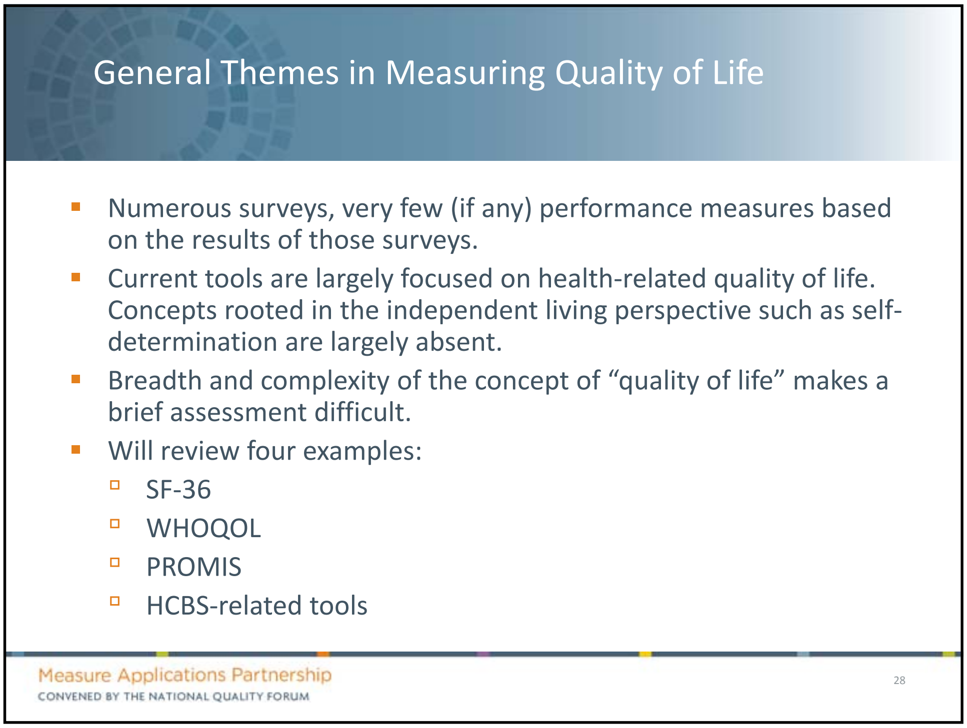
# Current Approaches to Assessing Quality of Life



## Definitions of "Quality of Life"

**Considerable agreement exists that quality of life is multidimensional.**

- Dictionary:
  - *Noun.* Your personal satisfaction (or dissatisfaction) with the cultural or intellectual conditions under which you live (as distinct from material comfort)
- Generally, the concept:
  - integrates objective and subjective indicators, a broad range of life domains, and individual values
  - incorporates physical, material, social, emotional, and developmental aspects



## General Themes in Measuring Quality of Life

- Numerous surveys, very few (if any) performance measures based on the results of those surveys.
- Current tools are largely focused on health-related quality of life. Concepts rooted in the independent living perspective such as self-determination are largely absent.
- Breadth and complexity of the concept of "quality of life" makes a brief assessment difficult.
- Will review four examples:
  - SF-36
  - WHOQOL
  - PROMIS
  - HCBS-related tools

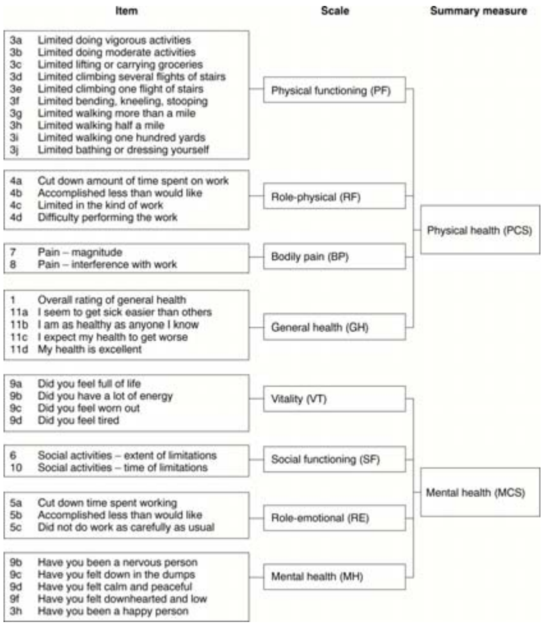
# RAND 36-Item Health Survey (SF-36)

- The SF-36® is a set of commonly used patient-reported quality-of-life measures.
- It was developed by RAND as part of the Medical Outcomes Study (MOS).
- It is a generic measure, as opposed to one that targets a specific age, disease, or treatment group.
- The 36 items were adapted from longer instruments. Shorter versions of the tool have also been developed (e.g., SF-12, SF-8) as has an open-access version (VR-12).
- It is available in many languages.
- Applications include:
  - Measuring health improvement or decline
  - Predicting medical expenses
  - Assessing treatment effectiveness
    - >600 randomized clinical trials
  - Comparing disease burden across populations
    - AHRQ has adopted the SF-12v2 for the Medical Expenditure Panel Study (MEPS)

Source: RAND Health. "36-Item Short Form Survey from the RAND Medical Outcomes Study
http://www.rand.org/health/surveys_tools/mos/mos_core_36item.html and http://www.sf-36.org/

---

- SF tools consist of two components (physical health and mental health).
- They are calculated with a total of eight scaled scores.



| Item | Scale | Summary measure |
|---|---|---|
| 3a Limited doing vigorous activities<br>3b Limited doing moderate activities<br>3c Limited lifting or carrying groceries<br>3d Limited climbing several flights of stairs<br>3e Limited climbing one flight of stairs<br>3f Limited bending, kneeling, stooping<br>3g Limited walking more than a mile<br>3h Limited walking half a mile<br>3i Limited walking one hundred yards<br>3j Limited bathing or dressing yourself | Physical functioning (PF) | Physical health (PCS) |
| 4a Cut down amount of time spent on work<br>4b Accomplished less than would like<br>4c Limited in the kind of work<br>4d Difficulty performing the work | Role-physical (RF) | |
| 7 Pain – magnitude<br>8 Pain – interference with work | Bodily pain (BP) | |
| 1 Overall rating of general health<br>11a I seem to get sick easier than others<br>11b I am as healthy as anyone I know<br>11c I expect my health to get worse<br>11d My health is excellent | General health (GH) | |
| 9a Did you feel full of life<br>9b Did you have a lot of energy<br>9c Did you feel worn out<br>9d Did you feel tired | Vitality (VT) | Mental health (MCS) |
| 6 Social activities – extent of limitations<br>10 Social activities – time of limitations | Social functioning (SF) | |
| 5a Cut down time spent working<br>5b Accomplished less than would like<br>5c Did not do work as carefully as usual | Role-emotional (RE) | |
| 9b Have you been a nervous person<br>9c Have you felt down in the dumps<br>9d Have you felt calm and peaceful<br>9f Have you felt downhearted and low<br>3h Have you been a happy person | Mental health (MH) | |

http://www.rand.org/health/surveys_tools/mos/mos_core_36item_survey_print.html

## Limitations of the SF Tools

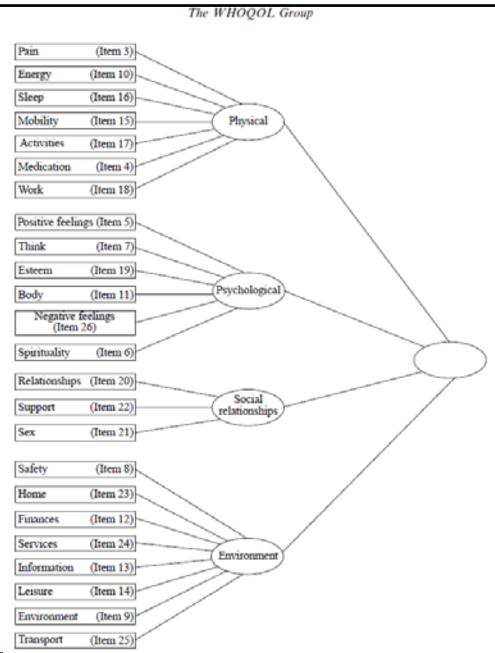
- SF tools do not assess signs and symptoms related to:
    - Sleeping patterns
    - Memory / concentration
    - Substance abuse
    - Sensory loss
    - Incontinence
    - Dizziness
    - … or many other topics of importance to people with needs for long-term services and supports
- The literature is mixed on the response and completion rates for certain populations (e.g., older adults)
- Usage restrictions on most current versions of the tool and scoring methodology

## World Health Organization WHO Quality of Life Instrument (WHOQOL)

- The World Health Organization's Quality of Life (WHOQOL) project began in 1991 to develop a cross-cultural quality of life assessment instrument.
- The original tool is a self-reported survey that contains 100 items.
    - A WHOQOL-BREF version contains 26 items.
    - An adaptation has been developed for people with HIV (WHOQOL-HIV) and an additional 32-item module assesses aspects of Spirituality, Religiousness and Personal Beliefs (WHOQOL-SRPB).
    - The WHO scales have been administered to more than 16,000 people in over 40 countries world-wide; they are available in more than 20 languages.

Sources: http://www.who.int/substance_abuse/research_tools/whoqolbref/en/ and http://depts.washington.edu/seaqol/docs/WHOQOL_Info.pdf

The WHOQOL-BREF assesses individuals’ perceptions in the context of their culture and value systems, and their personal goals, standards and concerns across four domains.



The WHOQOL Group. http://arno.uvt.nl/show.cgi?fid=75820

---

# Applications and Limitations of the WHOQOL

- The dimensions of quality of life in the WHO tools have a high level of international consensus.
- Some socio-demographic and health information is also obtained.
- Applications for the WHOQOL assessment include:
  - Medical practice
  - Assessing the effectiveness and relative merits of different treatments
  - Health services evaluation
  - Researching involving how disease impairs/impacts the subjective well-being of a person
- Known limitations of the assessment include:
  - Permission to use the WHOQOL is granted for each individual study.

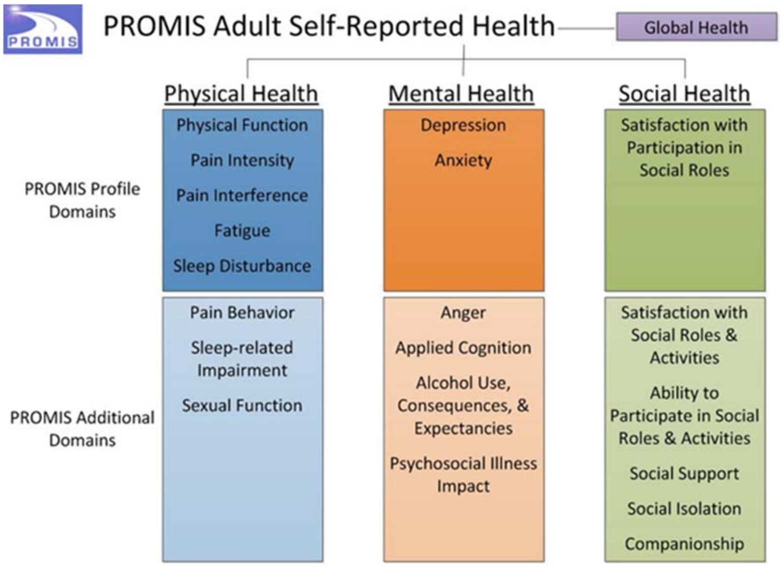
## Patient-Reported Outcomes Measurement Information System (PROMIS)

- In 2004, the National Institutes of Health (NIH) initiated a cooperative group called the Patient-Reported Outcomes Measurement Information System (PROMIS®)
- PROMIS provides a platform for self-reported measures of health, including symptoms, function and well-being. It offers:
  - Comparability—measures have been standardized so there are common domains and metrics across conditions
  - Reliability and Validity—all metrics have been rigorously reviewed and tested
  - Flexibility—it can be administered in a variety of ways
  - Inclusiveness—PROMIS encompasses all people, regardless of literacy, language, physical function or life course

Sources: http://www.nihpromis.org/

---

## PROMIS Adult Self-Reported Health

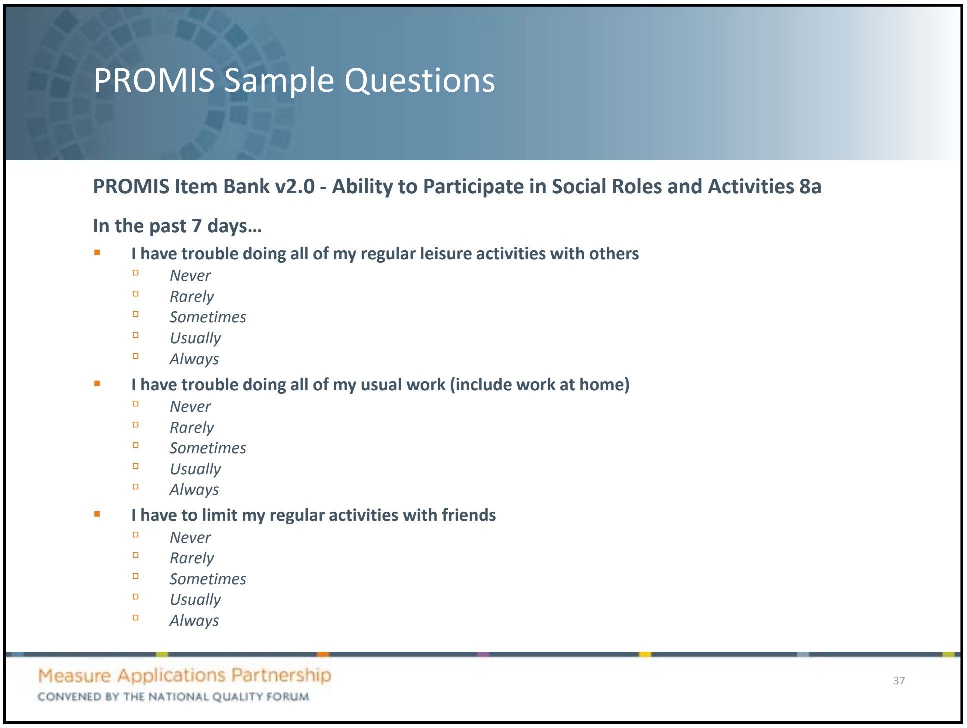


Global Health

| | Physical Health | Mental Health | Social Health |
|---|---|---|---|
| PROMIS Profile Domains | Physical Function<br>Pain Intensity<br>Pain Interference<br>Fatigue<br>Sleep Disturbance | Depression<br>Anxiety | Satisfaction with Participation in Social Roles |
| PROMIS Additional Domains | Pain Behavior<br>Sleep-related Impairment<br>Sexual Function | Anger<br>Applied Cognition<br>Alcohol Use, Consequences, & Expectancies<br>Psychosocial Illness Impact | Satisfaction with Social Roles & Activities<br>Ability to Participate in Social Roles & Activities<br>Social Support<br>Social Isolation<br>Companionship |

12/12/2012

http://www.nihpromis.org/measures/domainframework1#acc

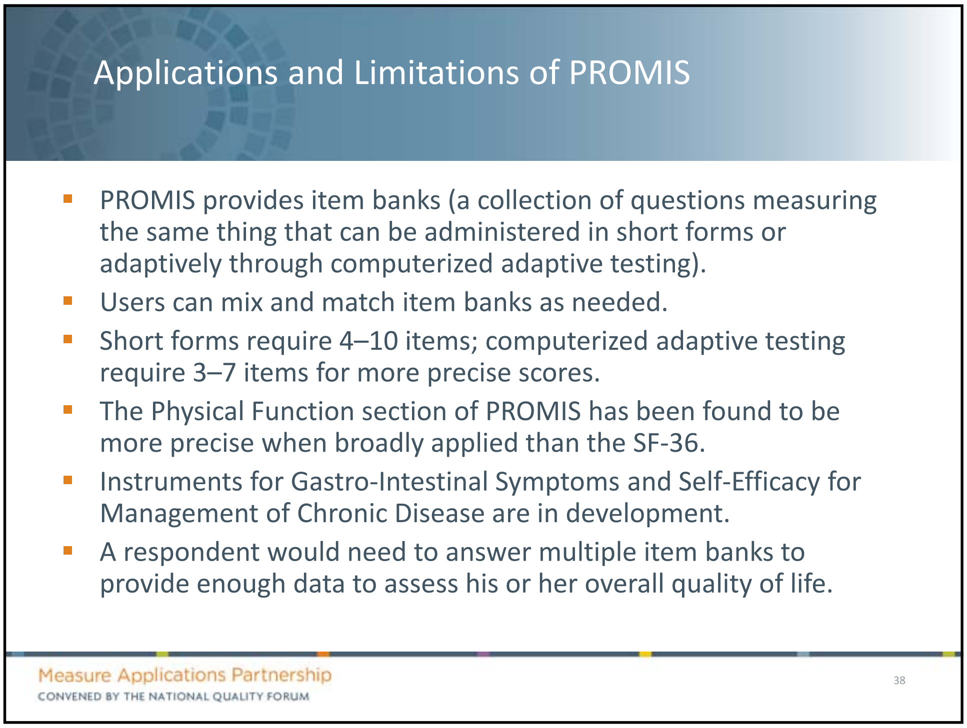


## PROMIS Sample Questions

**PROMIS Item Bank v2.0 - Ability to Participate in Social Roles and Activities 8a**

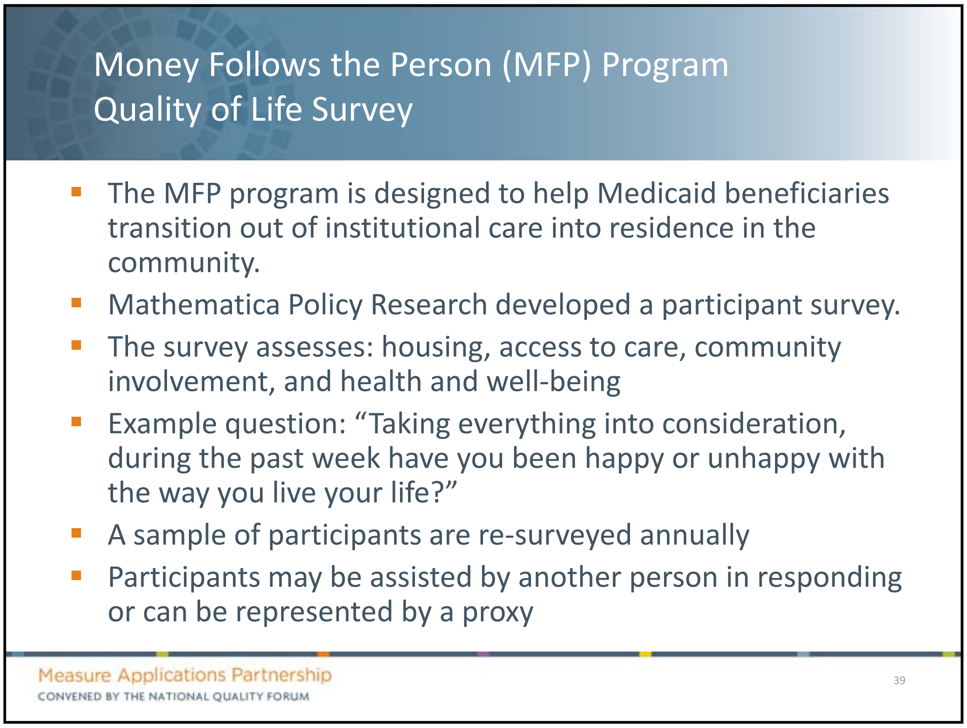
**In the past 7 days…**

- **I have trouble doing all of my regular leisure activities with others**
  - *Never*
  - *Rarely*
  - *Sometimes*
  - *Usually*
  - *Always*
- **I have trouble doing all of my usual work (include work at home)**
  - *Never*
  - *Rarely*
  - *Sometimes*
  - *Usually*
  - *Always*
- **I have to limit my regular activities with friends**
  - *Never*
  - *Rarely*
  - *Sometimes*
  - *Usually*
  - *Always*



## Applications and Limitations of PROMIS

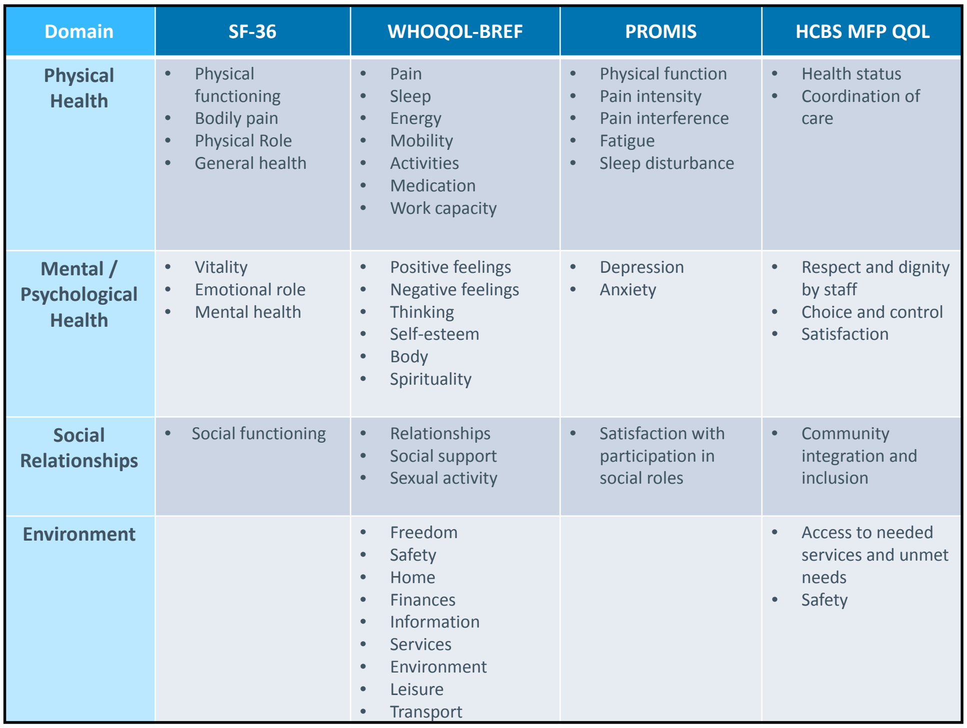
- PROMIS provides item banks (a collection of questions measuring the same thing that can be administered in short forms or adaptively through computerized adaptive testing).
- Users can mix and match item banks as needed.
- Short forms require 4–10 items; computerized adaptive testing require 3–7 items for more precise scores.
- The Physical Function section of PROMIS has been found to be more precise when broadly applied than the SF-36.
- Instruments for Gastro-Intestinal Symptoms and Self-Efficacy for Management of Chronic Disease are in development.
- A respondent would need to answer multiple item banks to provide enough data to assess his or her overall quality of life.



## Money Follows the Person (MFP) Program Quality of Life Survey

- The MFP program is designed to help Medicaid beneficiaries transition out of institutional care into residence in the community.
- Mathematica Policy Research developed a participant survey.
- The survey assesses: housing, access to care, community involvement, and health and well-being
- Example question: “Taking everything into consideration, during the past week have you been happy or unhappy with the way you live your life?”
- A sample of participants are re-surveyed annually
- Participants may be assisted by another person in responding or can be represented by a proxy

Measure Applications Partnership
CONVENED BY THE NATIONAL QUALITY FORUM



| Domain | SF-36 | WHOQOL-BREF | PROMIS | HCBS MFP QOL |
|---|---|---|---|---|
| **Physical Health** | • Physical functioning<br>• Bodily pain<br>• Physical Role<br>• General health | • Pain<br>• Sleep<br>• Energy<br>• Mobility<br>• Activities<br>• Medication<br>• Work capacity | • Physical function<br>• Pain intensity<br>• Pain interference<br>• Fatigue<br>• Sleep disturbance | • Health status<br>• Coordination of care |
| **Mental / Psychological Health** | • Vitality<br>• Emotional role<br>• Mental health | • Positive feelings<br>• Negative feelings<br>• Thinking<br>• Self-esteem<br>• Body<br>• Spirituality | • Depression<br>• Anxiety | • Respect and dignity by staff<br>• Choice and control<br>• Satisfaction |
| **Social Relationships** | • Social functioning | • Relationships<br>• Social support<br>• Sexual activity | • Satisfaction with participation in social roles | • Community integration and inclusion |
| **Environment** | | • Freedom<br>• Safety<br>• Home<br>• Finances<br>• Information<br>• Services<br>• Environment<br>• Leisure<br>• Transport | | • Access to needed services and unmet needs<br>• Safety |

## *Questions?*

## *Opportunity for Public Comment*

## Workgroup Perspectives: Challenges and Opportunities in Measuring Quality of Life

- Jamie Kendall, Administration for Community Living
- Anne Cohen, Subject Matter Expert: Disability
- Adam Burrows, National PACE Association
- Rich Bringewatt, SNP Alliance

- "Quality of life" is a broad concept; what are the necessary components to evaluate within the dual eligible beneficiary population?
- What methods should be used to measure quality of life within the dual eligible beneficiary population?
- What do you view as the primary barriers to measuring quality of life?
- What solutions do you propose to address those barriers?

---

# Quality of Life

## Measure Applications Partnership
## Dual Eligible Beneficiaries Workgroup Web Meeting

## November 21, 2013

Jamie Kendall, Administration for Community Living
Anne Cohen, MAP Disability Subject Matter Expert, Disability and Health Policy Consultant

# Domains

- culture, including language and health literacy
- housing
- family and friends
- employment
- community integration
- transportation
- behavioral health
- recreation
- vocational training
- relationship building
- Other choices

# Person Centered Planning Key

- choice and control
- availability and access
- dignity of risk

**Council on Quality and Leadership efforts:**

Measures the person outcomes (quality of life)

Measures the LTSS in place (services)

CQL:

Each outcome includes:

- Introduction and intent of outcome
- Values underlying outcome
- Principles for organizations
- Suggested questions for the person
- Suggested questions about individualized supports
- Outcome decision making questions
- Individualized support decision making questions
- Additional considerations

---

**National Core Indicators**

Document s:

- the effect of services on the day-to-day lives of

the people who receive them

- the experience of program participants
- Manage service delivery and improve policy and practice
- Track key performance goals and outcomes
- Assess the impact of regulatory activities on individual

Experience

- Respond to the demands of consumers and families for

information on system responsiveness

- Assess the impact of financial actions

Measures:

- Individual characteristics of people receiving services and support
- The locations where people live
- The activities they engage in during the day including
- whether they are working
- The nature of their experiences with the supports that
- they receive (e.g., with case managers,
- ability to make choices
- The context of their lives – friends,
- community involvement, safety
- Health and well-being, access to healthcare

## “Why Did Someone Like You Go SKYDIVING!?” Case Study in Dignity of Risk and the Response from the Health Care System

- Man with CP has a masters degree in conflict resolution, works as a Middle East Analyst. Full social life, lives on his own, performs his own tasks of daily living without assistance, drives his own car, exercises regularly, does not smoke or drink, eats healthy, does not receiving any government assistance.
- He skydives with 120 mile winds, a gust of wind partially deflated his parachute 20 feet off the ground. He came down on his right knee but walked off the field unassisted. Later goes to the ER with a swollen knee.

## Dignity of Risk
## The jump went great, then I went to the ER

**ER staff makes insensitive comments:**

- “Why are you using those two canes?”
- “Why would you do such a thing? We’re healthy over here and we’d never do such a thing.”
- “How did you drive yourself?”
- “Man, you’re a champ! I wouldn’t have the guts to do that and I don’t think anyone else would either.”

**The gentleman has to explain**

- “I use one cane because I have CP and the other because I hurt my knee skydiving.”
- “I am healthy, you should try skydiving sometime it’s amazing.”
- “I drove myself using hand controls”

**Multiple ER staff visit patient to see “man with CP who can drive himself using his hands and jumps from planes”**

# Questions to Ponder

**Quality of Life**

- Is quality of life a public health measurement domain or is it a health care quality issue, or both?

- Health care quality measurement considerations:
    - Cultural competency training, including needs of specific populations and knowledge of community resources, accessing community based LTC
    - Measurement domains associated with dignity of risk, like STD screening
    - Mining the sources of data that already exist (CMS duals demos, cash and counseling demonstrations, NCI, CQL, other)

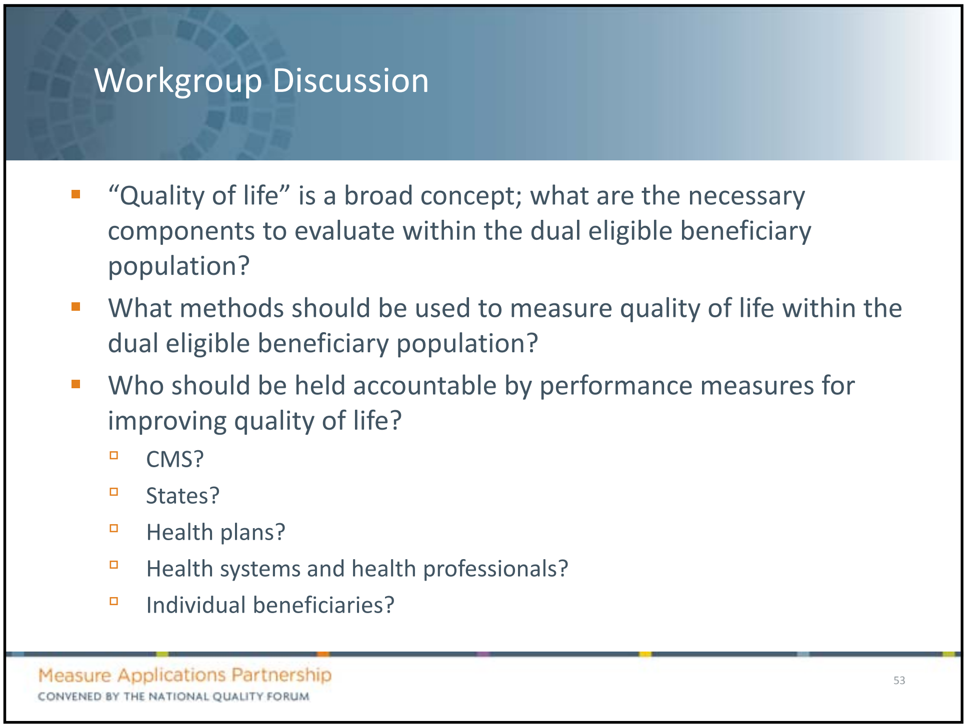
**Person Centered Measurements**

- How do we ensure that people and their individual preferences are at the center of our quality measure process?

- How do we address comprehensive families of measures?

# Appendices

- CQL data
- NCI data

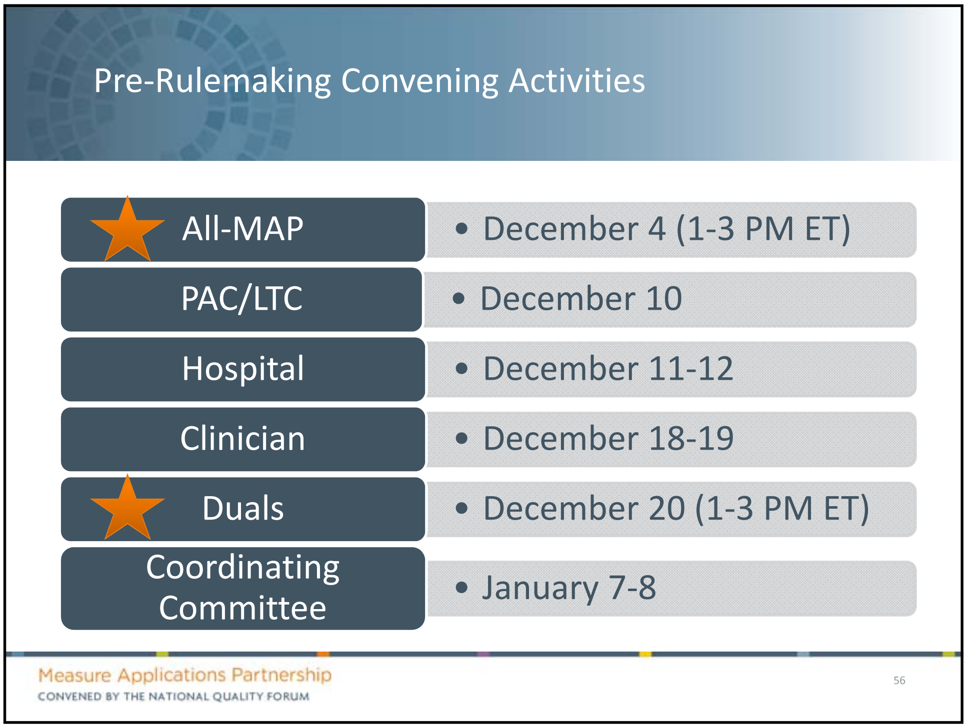


# Workgroup Discussion

- “Quality of life” is a broad concept; what are the necessary components to evaluate within the dual eligible beneficiary population?
- What methods should be used to measure quality of life within the dual eligible beneficiary population?
- Who should be held accountable by performance measures for improving quality of life?
  - CMS?
  - States?
  - Health plans?
  - Health systems and health professionals?
  - Individual beneficiaries?

# *Opportunity for Public Comment*

# *Wrap Up*



## Pre-Rulemaking Convening Activities

| | |
|---|---|
| ★ All-MAP | December 4 (1-3 PM ET) |
| PAC/LTC | December 10 |
| Hospital | December 11-12 |
| Clinician | December 18-19 |
| ★ Duals | December 20 (1-3 PM ET) |
| Coordinating Committee | January 7-8 |

## 2013/2014 Pre-Rulemaking Liaisons

| Clinician Workgroup | PAC/LTC Workgroup | Hospital Workgroup |
|---|---|---|
| Ruth Perry | Gwendolen Buhr | Jennifer Sayles |

- Liaisons will participate in other setting-specific MAP meetings to represent the perspective of vulnerable beneficiaries.
- Liaisons will report back to this workgroup on the December 20th web meeting.
- Alice will represent the workgroup at the MAP Coordinating Committee.

## Important Dates

- **December 2-10:** Early commenting on measures under consideration
- **December 20, 1:00-3:00 PM:** MAP Duals Workgroup Web Meeting for Pre-Rulemaking
- **January:** Public Comment on Draft Report on Pre-Rulemaking Input
- **January/February:** Workgroup Review and Public Comment on Draft Report on Dual Eligible Beneficiaries
  - High-need subgroups
  - Family of measures
  - Measurement of quality of life
- **February 1:** MAP 2013/2014 Pre-Rulemaking Recommendations Due to HHS
- **February 28:** MAP Report on Dual Eligible Beneficiaries Due to HHS

## Upcoming MAP Nominations

**Don't Forget:**

The annual 30-day call for nominations for new and renewing MAP members will begin in January

*Details will be available through the NQF website.*

---

***Thank you***

# Patient Reported Outcomes (PROs) in Performance Measurement

January 10, 2013

NATIONAL QUALITY FORUM

# Contents

# Patient Reported Outcomes (PROs) in Performance Measurement

## Introduction

### US Healthcare: Performance Improvement and Accountability

Widespread variation in the quality of healthcare in the United States is well documented.[1,2,3,4,5,6] Although many laudable examples can be identified across the country where safe, effective, affordable healthcare and long-term support services are consistently provided, serious gaps persist. Coupled with the need to constrain escalating costs of healthcare—threatening the livelihoods of individuals and families and the overall national economy—is the need to improve performance and hold providers accountable. The Patient Protection and Affordable Care Act of 2010 (hereafter, ACA) has several provisions targeting this challenge. One mandates creation of a National Quality Strategy (NQS) to serve as a blueprint to *improve the delivery of healthcare services, patient health outcomes, and population health.*[7] Released in March 2011 and updated yearly, the NQS identifies three overarching aims of better care, healthy people and communities, and affordable care; it also spells out six priority areas for collective action to drive toward a high-value health system (health and well-being, prevention and treatment of leading causes of mortality, person- and family-centered care, patient safety, effective communication and care coordination, and affordable care).[8,9]

### Achieving Performance Improvement & Accountability through Patient-Reported Outcomes

Patient and family engagement is increasingly acknowledged as a key component of a comprehensive strategy, (along with performance improvement and accountability), to achieve a high quality, affordable health system. Emerging evidence affirms that patients who are engaged in their care tend to experience better outcomes[10] and choose less costly but effective interventions, such as physical therapy for low back pain, after participating in a process of shared decisionmaking.[11] Promising approaches to involve patients and their families at multiple levels are being implemented across the country. Such activities include consumers serving on governance boards at hospitals and contributing to system and practice redesign to make care safer and more patient-centric.[12,13]

Historically, with the exception of collecting feedback on satisfaction or experience with care, patients remain an untapped resource in assessing the quality of healthcare and of long-term support services. Patients are a valuable and, arguably, the authoritative source of information on outcomes beyond experience with care. These include health-related quality of life, functional status, symptom and symptom burden, and health behaviors. For example, in the case of long-term support services for persons with disabilities, asking them about  outcomes they value, such as increased communication and self-help skills and improved social interactions is crucial. Hence, two critical steps are to engage patients by building capacity and infrastructure to capture patient-reported outcomes routinely and then to use these data to develop performance measures to allow for accurate appraisals of quality and efficiency.

## NQF Role in Promoting Accountability and Performance Improvement

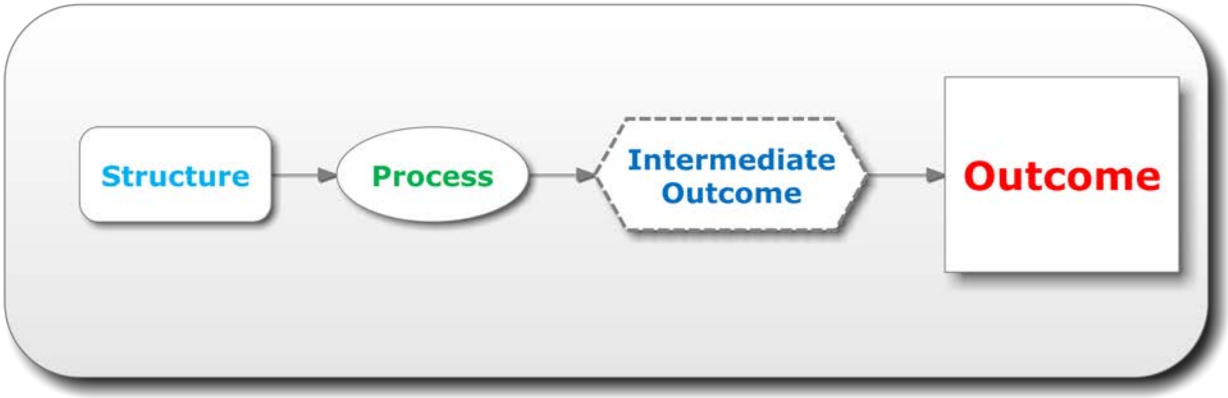
Valid, reliable measures are foundational for evaluating and monitoring performance and fostering accountability. The National Quality Forum (NQF) is a voluntary consensus standards-setting organization as defined by the National Technology Transfer and Advancement Act.[14] In this role, NQF endorses performance measures as consensus standards to assess the quality of healthcare for use in accountability applications such as public reporting and payment as well as performance improvement. NQF is a neutral evaluator of performance measures, not a measure developer. NQF convenes diverse stakeholders to evaluate measures based on well-vetted, widely accepted criteria.

The field of performance measurement is evolving to meet the demands of increased accountability to improve outcomes in both quality and costs. The direction for NQF-endorsed performance measures includes:

- driving toward higher performance reflected in outcome measures rather than in basic processes such as performing an assessment;
- measuring disparities;
- shifting toward composite measures that summarize multiple aspects of care;
- harmonizing measures across sites and providers; and
- conducting measurement across longitudinal, patient-focused episodes including outcome measures, process measures with direct evidence of impact on desired outcomes; appropriateness measures; and cost/resource use measures coupled with quality measures, including overuse.

Figure 1 depicts the relationship among structure, process, and outcome. NQF prefers to endorse performance measures of health outcomes that are linked directly to evidence-based processes or structures, or of outcomes of substantial importance with a plausible link to healthcare processes. Next in the preferred hierarchy are measures of intermediate outcomes and processes closely linked to desired outcomes. Measures of processes that are distal to desired outcomes (e.g., assess a clinical parameter) and those that are satisfied by a "checkbox" are considered to have the least impact on the goal of improving healthcare and health.

**Figure 1. Structure-Process-Outcome**



### *Patient-Reported Outcomes Tools & Performance Measures*

Patient-reported outcomes (PROs) are defined as "any report of the status of a patient's (or person's) health condition, health behavior, or experience with healthcare that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else."[15] "PRO" has become an international term of art; the word "patient" is intended to be inclusive of all persons, including patients, families, caregivers, and consumers more broadly. It is intended as well to cover all persons receiving support services, such as those with disabilities. Key PRO domains include:

- Health-related quality of life (including functional status);
- Symptoms and symptom burden (e.g. pain, fatigue);
- Experience with care; and
- Health behaviors (e.g., smoking, diet, exercise).

Various tools (e.g., instruments, scales, single-item measures) that enable researchers, administrators, or others to assess patient-reported health status for physical, mental, and social well-being are referred to as PRO measures (PROMs). In order to include PROs more systematically as an essential component of assessing the quality of care or services provided, and as part of accountability programs such a value-based purchasing or public reporting, it is necessary to distinguish between PROMs (i.e., tools) and aggregate-level performance measures.

A PRO-based performance measure (PRO-PM) is based on PRO data aggregated for an entity deemed as accountable for the quality of care or services delivered. Such entities can include (but would not be limited to) long-term support services providers, hospitals, physician practices, or accountable care organizations (ACOs). NQF endorses PRO-PMs for purposes of performance improvement and accountability; NQF does not endorse the PROMs alone. However, the specific PROM(s) used in a PRO-PM will be identified in the detailed measure specifications to ensure standardization and comparability of performance results. Table 1 illustrates the distinctions among PRO, PROM, and PRO-PM. Full definitions are in the glossary (see Appendix A).

**Table 1. Distinctions among PRO, PROM, and PRO-PM: Two Examples**

| Concept | Patients With Clinical Depression | Persons with Intellectual or Developmental Disabilities |
|---|---|---|
| **PRO** (patient-reported outcome) | Symptom: depression | Functional Status-Role: employment |
| **PROM** (instrument, tool, single-item measure) | PHQ-9©, a standardized *tool* to assess depression | Single-item measure on National Core Indicators Consumer Survey: *Do you have a job in the community?* |
| **PRO-PM** (PRO-based performance measure) | Percentage of patients with diagnosis of major depression or dysthymia and initial PHQ-9 score >9 with a follow-up PHQ-9 score <5 at 6 months (NQF #0711) | The proportion of people with intellectual or developmental disabilities who have a job in the community |

## PRO-PMs Applications: Benefits and Challenges

Interest and appreciation of the value of using PRO-PMs in performance measurement as part of the broader accountability and performance improvement landscape are mounting. To accelerate the adoption of PRO-PMs that can be used for these purposes several underlying issues must be addressed, which will require collaborative and collective effort across multiple stakeholder groups including providers, consumers, purchasers, measure developers, researchers, and others. First, PROMs have not been widely adopted for clinical use outside research settings in the United States; for that reason, they may be unfamiliar to many health professionals, payers, and provider institutions. Therefore, steps need to be taken to raise awareness among these stakeholders of the benefits of using PROMs to engage patients in their care improve outcomes. Second, there are several method-related challenges such as aggregating patient data on PROMs to measure performance at multiple levels of analysis (e.g., individual, group practice, organization) and use of proxy respondents. Therefore, more research is needed on best practices in this area.

### *Approach*

To begin to address these complex issues, NQF, with funding from the Department of Health and Human Services, is conducting the *PROs in Performance Measurement* project. The project goals are to:

- Identify key characteristics for selecting PROMs to be used in PRO-PMs;
- Identify any unique considerations for evaluating PRO-PMs for NQF endorsement and use in accountability or performance improvement applications; and
- Lay out the pathway to move from PROs to NQF-endorsed PRO-PMs.

NQF designed this project to bring together a diverse set of stakeholders (see Appendix B) who could facilitate the groundwork for developing, testing, endorsing and implementing PRO-PMs. Key steps in the project were to convene two workshops with an expert panel and to commission two papers. The papers focused on issues about methods and served as background for the workshops. The first paper focused on selecting PROMs for use in performance measurement and the second on the reliability and validity of PRO-PMs.[16,17]

National and international examples of successful experiences are encouraging. At the workshop, participants obtained valuable insights about approaches to data collection and aggregation and practical pointers about implementation (e.g., getting buy-in from providers). At the first workshop, colleagues from the Dartmouth Spine Institute and Massachusetts General Hospital presented information about their experiences using PROMs in patient care and performance improvement (available here). At the second workshop, representatives from the Centers for Medicare & Medicaid Services Health Outcomes Survey, England's National Health Service PROMs, and Sweden's national quality registers presented on their initiatives to report PRO-PMs publicly (available here). These discussions informed the recommendations found later in this report. Additionally, a large body of knowledge is available about PRO-PMs in the domain of experience with care (e.g., performance measures based on CAHPS®).

This report captures the insights from this effort to date and provides recommendations to move the field of performance measurement forward. The remaining sections of this report cover guiding principles, a detailed pathway from PROs to PRO-PMs, key implications and recommendations related to NQF endorsement criteria, and future directions.

## Guiding Principles

During the first workshop, members of the Expert Panel conceptualized “guiding principles” for selecting PROMs in the context of performance measurement: psychometric soundness, person-centered, meaningful, amenable to change, and implementable. They are not NQF endorsement criteria per se, but they serve as key constructs for recommendations on the pathway from PRO to PRO-PM. Measure developers should also take these principles into account in preparing submissions and documentation for NQF consideration for endorsement.

The guiding principles, described below, place the patient front and foremost. They underpin the thinking that shaped the pathway from PROs to PRO-PMs discussed in the next section of this report. The word “patient” is often used as shorthand to comprise patients, families, caregivers, and consumers more broadly. We also use this term to include persons receiving support services, such as those with disabilities. NQF must ensure that the emerging portfolio of PRO-PMs addresses a range of healthcare services that extend beyond the walls of a particular clinical setting of care.

### Psychometric Soundness

Workshop participants agreed on several characteristics that should be considered in selecting PROMs for use in PRO-PMs. These are listed in Box 1 and are derived from the first commissioned paper. Appendix C provides the expanded explanations for these properties of the instruments or tools used to measure PROs. Of these, reliability and validity are considered baseline requirements for selecting a PROM to use in a performance measure. The remaining principles below presume that reliability and validity of a selected PROM was adequately demonstrated.

Box 1. Characteristics for Selecting PROMs Identified in Commissioned Paper

1. Conceptual and Measurement Model Documented
2. Reliability
    - 2a. *Internal consistency (multi-item scales)*
    - 2b. *Reproducibility (stability over time)*
3. Validity
    - *3a. Content Validity*
    - *3b. Construct and Criterion-related Validity*
    - *3c. Responsiveness*
4. Interpretability of Scores
5. Burden
6. Alternatives modes and methods of administration
7. Cultural and language adaptations
8. Electronic health record (EHR) capability

### Person-Centered

“Person-centeredness” was the overarching theme that arose from the workshop discussions. In this context, using PROMs is viewed as an important step toward engaging patients, health professionals, and other entities in creating a person-centered health system. Workshop participants also identified the opportunity for PROMs to facilitate shared decisionmaking (SDM), another strategy for engaging patients. SDM is defined as *a collaborative process that allows patients and their providers to make healthcare decisions together, taking into account the best scientific evidence available, as well as the patient’s values and preferences*.[18] For SDM, clinicians and other healthcare staff can use the instrument, scale, or single-item measure (PROM) to engage patients in their own preferred self-management and goal attainment by identifying outcomes important to them and tracking change over time. An important caveat to this discussion is that not all patients want to engage in formal SDM activities. Therefore, although contributing to SDM efforts is desirable, not all PROMs need to enable SDM. Measures of decision quality, defined as *the match between the chosen option and the features* *that matter most to the informed patient,* fall under the PRO domain of experience with care and take into account whether patients are informed with the best available evidence and there is concordance between what matters most to them and the treatments they receive.[19,20]

As a final consideration of person-centeredness, as patients become more engaged in their care by providing systematic feedback on outcomes such as their functional or health status, the flow of information between clinicians and patients must be bi-directional. This may mean that health professionals interpret PROM information back to their patients; it may mean that mechanisms are established to give patients their own information directly (displayed in easy-to-understand ways). With steps such as these, respondents to PROMs can benefit from seeing results in a timely way, and this type of service can balance any perceived burdens they may feel about completing data-collection activities. Although these considerations may not affect NQF endorsement efforts directly, the Expert Panel wished to emphasize that PROMs that can be used in this manner are desirable.

## Meaningful

Closely intertwined with person-centeredness is the concept of "meaningfulness." Meaningfulness encompasses the relevance and degree of importance of the concepts measured by the PROM from the perspective of patients, their families, and caregivers as well as clinicians and other health professionals who serve them. Among the concepts that PROMs would ideally capture are the following: the impact of health-related quality of life (including functional status); symptom and symptom burden; experience with care and satisfaction with the services; perceived utility of the services for achieving personal goals; or health behaviors. As suggested above, the focus comprises both "traditional" healthcare services, broadly defined, and support services for persons with disabilities.

Workshop participants debated how best to demonstrate that stakeholders think a particular PROM is meaningful. The following framework, coined as the three "Cs," are three aspects of meaningfulness on which to seek patient input and can serve as a starting point for thinking about how to operationalize this construct:

- Conceptual: engaging people in the dialogue about what matters most to them to define the concepts that PROs should cover. This upstream interaction is critical to meet a threshold consideration that what is being measured is important and meaningful to patients.

- Contextual: learning how individuals use the information derived from either a PROM or a PRO-PM. Individuals here are defined very broadly to include not just patients (however construed for the application at hand) but also clinicians, other health professionals, administrators, and perhaps even policymakers. For example, does such information facilitate their participation in managing their own healthcare? Does it help people to select a high-quality provider of healthcare or support services? Do such data contribute to the discourse on larger social issues such as achieving high-quality care at acceptable costs?

- Consequential: determining what happens when PRO-PM information is used in accountability applications (e.g., public reporting, value-based purchasing) or performance improvement. Performance data on PRO-PMs can have important consequences on the availability and receipt of quality services, the type of services, and their responsiveness to individuals' needs.

## Amenable to Change

Amenable to change refers to evidence that the outcome of interest is responsive to a specific healthcare or support service or intervention. The Panel noted that an outcome may be something a patient wants but the clinician or service provider cannot help to achieve the desired change. This concept applies to PROs as well as other types of outcomes. The reasoning is that outcome performance measures (including PRO-PMs) intended for both accountability and improvement should be supported by evidence that the providers being evaluated can influence the person's short- or long-term outcomes. Therefore, without such evidence, the performance measure would not be considered a valid indicator of quality of care.

However, a unique aspect of PROs (in contrast to other outcomes) is that patients must supply the data, which requires effort. As PROMs are collected more routinely at the point of care and are embedded into workflows, it becomes essential to ensure this information is of value to the patient and will be used in their care. Analogous to the collection of a blood sample to measure glucose concentration over time for diabetes (e.g., HbA1c), the PROM data should be relevant to the individual's care. Results should be shared and appropriate intervention taken (or not) based on the best available evidence and informed by patient's preferences and treatment goals. When collecting individual level data through the use of PROMs, special consideration must be given to the burden of data collection, which ideally will be offset by the patient's assessment of meaningfulness.

There was robust discussion among the workshop participants on this proposed guiding principle during which a strong counter argument was aired with respect to the idea that all PROs considered for purpose of accountability or performance measurement must be amenable to change. The rationale was that some outcomes are worth measuring even though they may not, at this point in time, be amenable to change by providers. For example, there are some outcomes (e.g., time to recovery) that are meaningful to patients that may not currently be considered modifiable, but nonetheless provide valuable information to patients and help them, working closely with their care provider, to make informed decisions. Additionally, if the outcome is deemed of high importance to patients, the process of measuring and reporting it could identify variation in performance and facilitate the spread of effective interventions. Although this disagreement was not resolved at the workshop, the point merits exploration.

## Implementable

The guiding principle that a PROM should be "implementable" acknowledges that many diverse factors affect their practical use in quality or accountability programs. Most of these factors relate to barriers to adopting such tools (PROMs) or collecting data and reporting on PRO-PMs in many practices, institutions, or other settings. Workshop participants raised many implementation issues. Although the examples were not exhaustive, workshop participants emphasized issues on the following list: administering PROMs in real-world situations; addressing literacy and health literacy of respondents; addressing cultural competency of clinicians and other service providers; dealing with the potential for unintended consequences related to patient selection; covering costs associated with using PROMS (especially those not available in the public domain); and adapting PROMs to computer-based platforms or other formats.

## Pathway from PRO to NQF-Endorsed PRO-PM

The pathway displayed in Figure 2, and described in detail below lays out the critical steps in developing a PRO-based performance measure suitable for endorsement by NQF and generating the evidence that it meets NQF criteria for endorsement. It begins with the conceptual basis for identifying a PRO for performance measurement; the pathway then proceeds through selecting a PROM and developing and testing a performance measure to achieving NQF endorsement of a PRO-PM and using the performance measure for accountability and performance improvement. This pathway describes how a PROM may form the basis of a PRO-PM that NQF could eventually endorse based on the NQF criteria.

The quality performance measurement enterprise includes multiple stakeholders who collaborate to develop performance measures, including methodologists and statisticians, as well as those receiving care and services, those whose performance will be measured, and those who will use performance results. In this discussion, the reference to developers includes all the participants in developing performance measures, not just formal measure developer organizations.

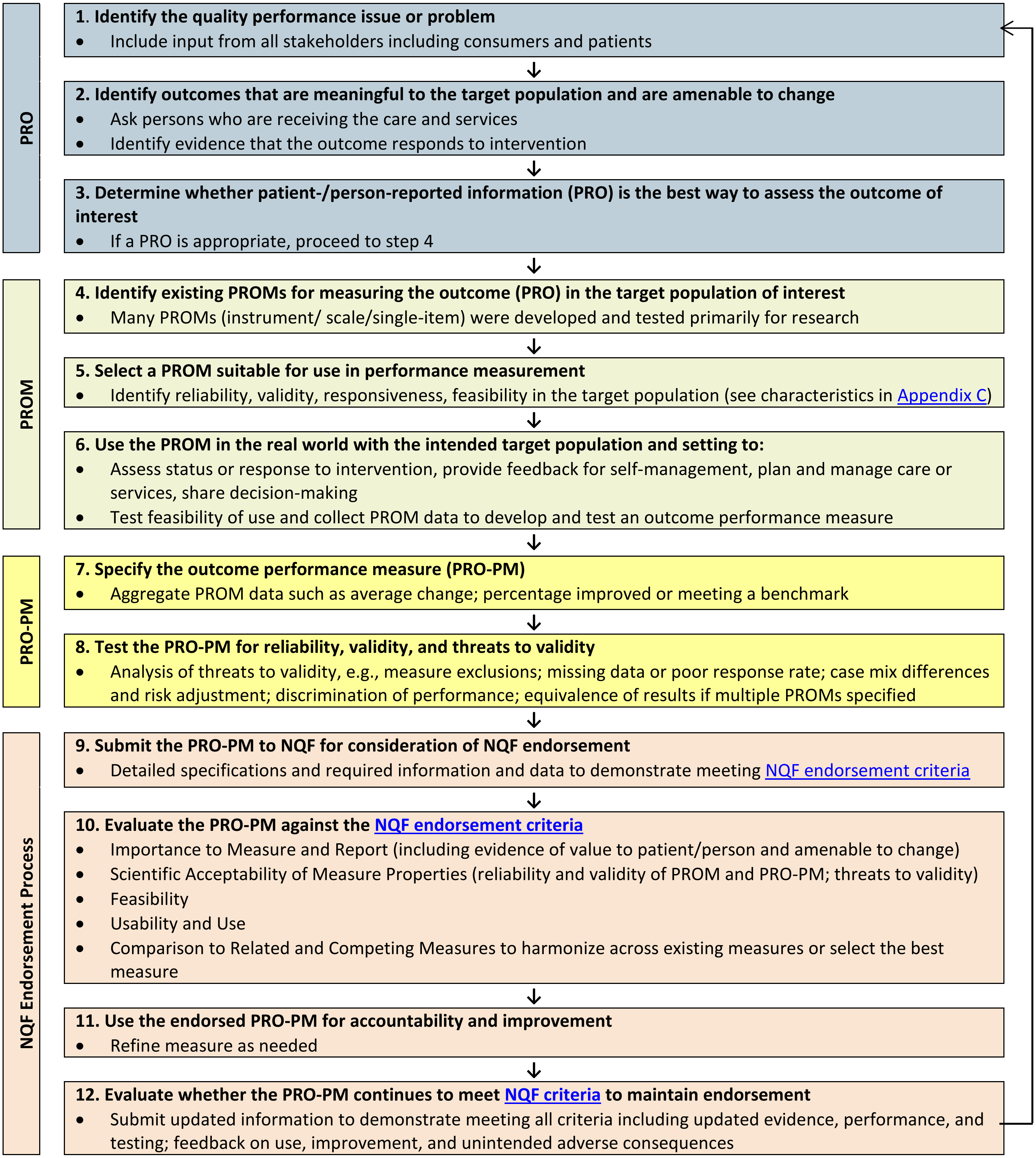
Although NQF is involved in the last section of the pathway, the earlier steps have implications for whether a performance measure will be suitable for NQF endorsement. Thus, they are intended to serve as a guide and best practices to help ensure that PRO-PMs will meet NQF criteria. For example, steps 1 and 2 in the pathway indicate that patients (as broadly defined as above) should be involved in identifying quality issues and outcomes that are meaningful to those receiving the healthcare and support services. If patients are involved at those steps, then developers will have amassed the information needed to demonstrate that the outcome is of value to patients. In the context of using this pathway leading to an NQF-endorsed performance measure, step 2 also suggests that developers identify outcomes with evidence that the outcome is responsive to intervention.

The steps shown in Figure 2 and described below are intended to help ensure that a proposed performance measure will meet NQF criteria for endorsement.

**Recommendation 1.**

Those developing PRO-PMs to be considered for NQF endorsement should follow the basic steps shown in the pathway in Figure 2. Doing so will help ensure that the eventual PRO-PM and its supporting documentation conform to NQF endorsement criteria.

**Figure 2. Pathway from PRO to NQF-endorsed PRO-PM**



**PRO**

**1. Identify the quality performance issue or problem**
- Include input from all stakeholders including consumers and patients

**2. Identify outcomes that are meaningful to the target population and are amenable to change**
- Ask persons who are receiving the care and services
- Identify evidence that the outcome responds to intervention

**3. Determine whether patient-/person-reported information (PRO) is the best way to assess the outcome of interest**
- If a PRO is appropriate, proceed to step 4

**PROM**

**4. Identify existing PROMs for measuring the outcome (PRO) in the target population of interest**
- Many PROMs (instrument/ scale/single-item) were developed and tested primarily for research

**5. Select a PROM suitable for use in performance measurement**
- Identify reliability, validity, responsiveness, feasibility in the target population (see characteristics in Appendix C)

**6. Use the PROM in the real world with the intended target population and setting to:**
- Assess status or response to intervention, provide feedback for self-management, plan and manage care or services, share decision-making
- Test feasibility of use and collect PROM data to develop and test an outcome performance measure

**PRO-PM**

**7. Specify the outcome performance measure (PRO-PM)**
- Aggregate PROM data such as average change; percentage improved or meeting a benchmark

**8. Test the PRO-PM for reliability, validity, and threats to validity**
- Analysis of threats to validity, e.g., measure exclusions; missing data or poor response rate; case mix differences and risk adjustment; discrimination of performance; equivalence of results if multiple PROMs specified

**NQF Endorsement Process**

**9. Submit the PRO-PM to NQF for consideration of NQF endorsement**
- Detailed specifications and required information and data to demonstrate meeting NQF endorsement criteria

**10. Evaluate the PRO-PM against the NQF endorsement criteria**
- Importance to Measure and Report (including evidence of value to patient/person and amenable to change)
- Scientific Acceptability of Measure Properties (reliability and validity of PROM and PRO-PM; threats to validity)
- Feasibility
- Usability and Use
- Comparison to Related and Competing Measures to harmonize across existing measures or select the best measure

**11. Use the endorsed PRO-PM for accountability and improvement**
- Refine measure as needed

**12. Evaluate whether the PRO-PM continues to meet NQF criteria to maintain endorsement**
- Submit updated information to demonstrate meeting all criteria including updated evidence, performance, and testing; feedback on use, improvement, and unintended adverse consequences

## Pathway Section Related to the PRO

The pathway begins with the conceptual basis for identifying a PRO for performance measurement.

### *1. Identify the quality performance issue or problem.*

Before developers devote resources to performance measurement, they need a clear understanding of the quality performance issue or problem related to healthcare or support services for a target population. Such understanding will direct the focus and establish the need for a performance measure. Input from all stakeholders including the recipients of the care and services, providers whose performance will be measured, payers, purchasers, and policymakers is critical to identifying priorities for performance measurement.

### *2. Identify outcomes that are meaningful to the target population and are amenable to change.*

After developers articulate the quality performance issue, they should identify the specific outcomes that are valued and meaningful to the target population in the context of a specific healthcare or support service. That is, the people receiving the services should be asked for their input. At this stage, all relevant desired outcomes should be identified even if they might not be assessed through patient-reported data.

As noted previously, the Expert Panel discussed focusing performance measures on outcomes that are responsive to intervention by healthcare and support service providers. The reason for this is twofold: 1) so patients are only asked to provide PROM data that is directly applicable to their care and treatment, and 2) so that providers' performance is measured on outcomes influenced by the care they provide. While there may be reasons to measure performance on important outcomes without such evidence, outcomes with evidence that they are influenced by at least one structure, process, intervention, or service should be considered as a starting point to garner broad-based support.

### *3. Determine whether patient- or person-reported information is the best way to assess the outcome of interest.*

Patient- or person-reported data are not necessarily the best way to assess every desired outcome identified in the prior step. The domains of health-related quality of life including functional status, symptoms and symptom burden, and health behaviors are outcomes for which individuals receiving healthcare and support services may be the best or only source of information. However, other meaningful outcomes such as survival (or mortality) and hospital readmission could be assessed using other data sources. If a PRO is the best approach to assess quality, then the pathway continues with the next section (steps 4-6); if not, then usual processes for specifying and testing measures and NQF review and endorsement, similar to steps 7-12 are followed.

## Pathway Section Related to the PROM

Given that one or more PROs are identified in the above steps, the pathway addresses the steps that organizations should take to select a PROM suitable to use in a performance measure.

### *4. Identify existing PROMs for measuring the outcome (PRO) in the target population of interest.*

As many PROMs already exist, developers should use various strategies (e.g., literature searches, PROMIS, web searches, outreach to experts in the field) to search for and identify PROMs that measure the outcome of interest in the target population. PROMs that were developed years ago may not have benefited from patient input; therefore, including patients in selecting PROMs to be used in performance measures is important.

### *5. Select a PROM suitable for use in performance measurement.*

The scientific (psychometric) characteristics that organizations should examine in selecting a PROM for performance measurement were summarized above and appear in detail in Appendix C. Of great importance is that PROMs be reliable, valid, and responsive in the target population. If no PROM for the target population seems to be suitable for use in a performance measure, then a developer or research group should test one or more PROMs in the target population or develop and test a wholly new PROM before a performance measure can be developed. The commissioned paper on methods issues related to PROMs is a resource on considerations for selecting PROMs (available here).[16]

### *6. Use the PROM in the real world with the intended target population and in the intended setting.*

The Expert Panel agreed that developers should collaborate with providers to use PROMs with the target population and in the settings for which performance measures are proposed before developing a PRO-PM. Many PROMs were developed for research studies and the resources and protocols for administering PROMs may not be realistic for broad scale implementation. A real-world application will identify feasibility issues related to administration, data capture, and workflow to use the PROM to assess individuals' responses to healthcare or support services, provide feedback for self-management, and (as desired) facilitate shared decisionmaking. At the first workshop, representatives from Dartmouth Spine Center and Partners Healthcare presented their experiences with using PROMs in clinical practice (available here).

Actual use of the PROM also generates the data needed to determine the best way to aggregate the PROM data in a performance measure and test the PRO-PM for reliability and validity. Widespread implementation is not a prerequisite for NQF endorsement; however, testing for reliability and validity and addressing risk adjustment are required. The data for such testing could come from settings that have already implemented the PROM, a pilot study, or a broader demonstration. An endorsed performance measure focused on the process of administering the PROM is not a necessary prerequisite and could divert resources and slow the endorsement of PRO-PMs. Performance measures focused on administering a PROM may not meet NQF criteria for endorsement and is discussed under recommendations related to the NQF evaluation criteria.

## Pathway Section Related to the PRO-PM

After the developer has selected the PROM and collaborated with providers who used it in practice to generate sufficient data for testing, the pathway addresses how developers should specify and test a PRO-PM.

### *7. Specify the outcome performance measure (PRO-PM).*

Developers specify how the outcome performance measure will be constructed. The metrics may be, for instance, an average change, the percentage of patients improved, or the percentage of respondents meeting a specific benchmark value. The performance measure needs to be fully specified including the specific PROM(s) used, guidance for administration, and rules for scoring; it should also describe the target population and any exclusions, give time frames for PROM administration and performance measurement, and outline any needed risk adjustment procedures.

### *8. Test the PRO-PM for reliability, validity, and threats to validity.*

Developers need to test the performance measure for reliability and validity. They explicitly need to address a variety of threats to validity or other technical issues. These include the need for risk adjustment or stratification and options for doing this, appropriateness of potential exclusions, and options for dealing with missing data. A further challenge is demonstrating equivalence of results when multiple PROMs are used.

Testing the PRO-PM is distinct from testing the PROM. Using a PROM with sound psychometric properties is necessary but not sufficient to assure a reliable and valid PRO-PM. The commissioned paper on methods issues for PRO-PMs provides a resource on considerations and approaches to examining or demonstrating reliability and validity of the performance measure (available here).[17]

## Pathway Section Related to the NQF Endorsement Process

The last section of the pathway focuses on the NQF endorsement process.

### *9. Submit the PRO-PM to NQF for consideration of NQF endorsement.*

The NQF endorsement process begins when developers submit a measure to NQF for consideration. Developers submit required information in NQF’s standard form so that all the information needed to evaluate the measure is available to reviewers.

### *10. Evaluate the PRO-PM against the NQF Endorsement Criteria.*

NQF evaluates measures against four main endorsement criteria listed here and described and discussed in more detail below.

1. Importance to Measure and Report
2. Scientific Acceptability of Measure Properties
3. Feasibility
4. Usability and Use

In addition, NQF has criteria and processes to address measure harmonization and selection of the best measure from among competing measures, which also would apply to PRO-PMs.

### *11. Use the endorsed PRO-PM for accountability and improvement.*

Once endorsed, NQF expects the measure to be used for accountability and performance improvement applications. Implementation of the performance measure should facilitate the goal of improvement and allow for measuring and tracking performance. Use of the performance measure provides data on performance to be examined for intended and unintended consequences.

In the case of PRO-PMs initially endorsed with testing based on limited PROM data, implementation of the PRO-PM could be phased. The initial emphasis would be on collecting the PROM data to expand testing and refine the measure before reporting performance on the outcome.

### *12. Evaluate whether the PRO-PM continues to meet NQF Criteria to maintain endorsement.*

NQF reviews each endorsed measure every three years to evaluate whether it continues to meet NQF criteria. In making its decision at this stage, NQF evaluates the measure on all criteria and considers information on actual use, improvement, and unintended adverse consequences. This information and the NQF endorsement maintenance decision also provide feedback to developers who are considering developing performance measures based on PROs.

# Key Implications and Recommendations Related to NQF Criteria

## Overview

The NQF endorsement criteria and guidance on evaluating all performance measures also apply to PRO-PMs. The four main endorsement criteria are: importance to measure and report, scientific acceptability of measure properties, feasibility, and usability and use. NQF committee members use the criteria to evaluate measures submitted for potential endorsement. When the performance measure meets the relevant criteria and NQF endorses a measure, it is considered suitable for purposes of accountability and performance improvement. Potential submitters (i.e., developers) need to be very familiar with the NQF criteria to be able assemble the required documentation as part of their submission.

PRO-PMs may, however, have unique aspects that warrant special consideration for measure evaluation and they are identified in this section. In addition, the exploration of PRO-PMs in this project highlighted some issues that also are relevant to other performance measures and those recommendations also are included in this section. The Expert Panel agreed that PRO-PMs should be held to the same criteria as other performance measures. Therefore, some of the panel's recommendations must be considered in the larger context of NQF endorsement criteria for all performance measures, specifically the recommendations related to the evidence criterion (3, 4, and 5). If the criteria are revised, the same standards would be applied to PRO-PMs as any other outcome performance measure.

Table 2 lists considerations for evaluation in the context of the main NQF endorsement criteria. The left column provides an abbreviated description of each criterion. The middle column identifies some considerations that PRO-PMS bring to light, but they are not unique to PRO-PMs. Several unique aspects about PRO-PMs are identified in the right column. This section provides recommendations and rationales for modifying the NQF criteria or guidance.

**Table 2. NQF Endorsement Criteria and their Application to PRO-PMs**

| Abbreviated NQF Endorsement Criteria | Considerations for Evaluating PRO-PMs that are relevant to other performance measures | Unique Considerations for Evaluating PRO-PMs |
|---|---|---|
| **Importance to Measure and Report**<br>a. High impact<br>b. Opportunity for improvement<br>c. Health outcome OR evidence-based process or structure of care | • Does evidence support that the outcome is responsive to intervention?<br>• When should the evidence exception be allowed for performance measures focused solely on conducting an assessment (e.g., administering a PROM, lab test)? | • Patients/persons must be involved in identifying PROs for performance measurement (person-centered; meaningful). |
| **Scientific Acceptability of Measure Properties**<br>a. Reliability<br>1. precise specifications<br>2. reliability testing for either data elements or performance measure score<br>b. Validity<br>1. specifications consistent with evidence<br>2. validity testing for either data elements or performance measure score<br>3. exclusions<br>4. risk adjustment<br>5. identify differences in performance<br>6. comparability of multiple data sources | • Data collection instruments (tools) should be identified (e.g., specific PROM instrument, scale, or single item).<br>• If multiple data sources (i.e., PROMs, methods, modes, languages) are used, then comparability or equivalency of performance scores should be demonstrated. | • Specifications should include standard methods, modes, languages of administration; whether (and how) proxy responses are allowed; standard sampling procedures; how missing data are handled; and calculation of response rates to be reported with the performance measure results.<br>• Reliability and validity should be demonstrated for both the data (PROM) and the PRO-PM performance measure score.<br>• Response rates can affect validity and should be addressed in testing.<br>• Differences in individuals' PROM values related to PROM instruments or methods, modes, and languages of administration need to be analyzed and potentially included in risk adjustment. |

| **Abbreviated NQF Endorsement Criteria** | **Considerations for Evaluating PRO-PMs that are relevant to other performance measures** | **Unique Considerations for Evaluating PRO-PMs** |
|---|---|---|
| **Feasibility**<br>a. Data generated and used in care delivery<br>b. Electronic data<br>c. Data collection strategy can be implemented | • The burdens of data collection, including those related to use of proprietary PROMs, are minimized and do not outweigh the benefit of performance measurement. | • The burden to respondents (people providing the PROM data) should be minimized (e.g., availability and accessibility enhanced by multiple languages, methods, modes).<br>• Infrastructure to collect PROM data and integrate into workflow and EHRs, as appropriate. |
| **Usability and Use**<br>a. Accountability and transparency<br>b. Improvement<br>c. Benefits outweigh unintended negative consequences | • Adequate demonstration of the criteria specified above supports usability and ultimately the use of a PRO-PM for accountability and performance improvement. | |

## Evidence that the PRO is of Value to the Target Population

**Recommendation 2.**

The NQF criterion or guidance for importance to measure and report should require evidence that the target population values the measured PRO and finds it meaningful.

Person-centeredness is a key principle for developing PRO-PMs. As shown in Figure 2, identifying outcomes of value to the target population is a critical early step in the pathway to endorse a PRO-PM. NQF's current criteria require evidence that the aspect of care being measured is of value to the patient when measures of experience with care are being evaluated. Experience with care is considered one type of patient-reported outcome; therefore, the requirement for having evidence of the value to the patient needs to be expanded to apply to all patient-reported outcomes.

## Evidence that the Measured PRO is Responsive to Intervention

**Recommendations 3-4.**

**3.** NQF should consider adding a criterion or guidance related to evidence to require identification of the causal pathway linking the relevant structures; (processes, interventions, or services); intermediate outcomes; and health outcomes.

**4.** NQF should consider applying the existing criterion and guidance regarding evidence for a process performance measure to health outcome performance measures, including PRO-PMs – i.e., *a systematic assessment and grading of the quantity, quality, and consistency of the body of empirical evidence* that at least one of the identified healthcare structures, processes, interventions, or services influences the outcome.

Currently the NQF measure submission asks developers to identify structure-process-outcome linkages, which serves as the introduction to the evidence section. It is appropriate for any type of performance measure and should serve as the basis for describing the proximity to desired health outcomes, identifying the appropriate evidence base for a performance measure, or identifying structures or processes that affect health outcomes.

Amenable to change was a key principle identified for developing PRO-PMs; however, the discussion and rationale extended to health outcome measures, in general. The Expert Panel suggested that evidence that the PRO or health outcome is responsive to intervention be required for NQF endorsement for all outcome performance measures. This represents a departure from NQF's current NQF guidance regarding evidence for performance measures of health outcomes and will require further examination by the CSAC and Board and a plan and timeline for implementation if this approach is recommended by these bodies.

For health outcome measures, NQF requires only a rationale linking the outcome to at least one healthcare structure, process, intervention, or service; it does not require submitting and evaluating information on systematic reviews of the empirical body of evidence, as required for other types of performance measures. NQF's position on evidence for health outcomes is based on the following reasoning:

- Health outcomes such as survival, physical or cognitive function, relief of symptoms, or prevention of morbidity are the reasons for seeking care and the goal of providing care. Therefore, these outcomes are central to measuring the performance of those rendering healthcare or support services.
- Health outcomes are often integrative. As such, they may reflect the influence of multiple clinicians and care processes and therefore are based on multiple bodies of evidence. Submitting information on multiple bodies of evidence could be burdensome and a disincentive for submitting outcome performance measures for NQF endorsement.
- Measuring health outcomes to identify variability in performance is a key driver to identifying strategies for improvement, even for outcomes previously thought not to be modifiable such as central line-associated bloodstream infections.

As discussed under the guiding principles, these same rationales were identified for PROs. However, the current environment in which penalties may be associated with performance measure scores has increased concern about using outcome performance measures for accountability. To mitigate that concern to some extent, the Expert Panel suggested focusing performance measurement on PROs that are meaningful to patients *and* with evidence that they are responsive to intervention. England and Sweden are leaders in the area of measuring PROs for performance measurement and appear to have taken this approach. England measures and reports performance on PROMs focused on specific surgical procedures to ameliorate problems with function and symptoms-hip and knee replacement and varicose vein surgery (access reports here). Sweden measures and reports performance on PROMs related to surgical procedure outcomes and complications (access report here). Sweden also reports performance on PROMs for a few medical conditions such as functioning three months after a patient has suffered a stroke and improvement after patients have started biological drug therapy for rheumatoid arthritis.

The Expert Panel acknowledged the trade-offs to a condition-specific approach. First, it excludes much of the population receiving healthcare and long-term support services. Second, even for a specific condition, limiting performance measurement to those who received only one possible intervention (e.g., surgery) does not provide a complete picture of performance related to the condition. A related question is whether to measure the PRO with generic or condition-specific PROMs. Condition-specific PROMs may be more responsive to change. However, generic measures offer more breadth, which is relevant, given that many patients have more than one condition. Using both generic and condition-specific PROMs affords the opportunity to better understand the benefits and drawbacks of both. These issues will need to be considered and revisited as the field gains experience with PRO-PMs.

### *Performance Measure Focused on Administering a PROM*

**Recommendation 5**

NQF should consider providing explicit guidance when a performance measure focused on collecting assessment data, including administering a PROM, meets the exception for the evidence criterion and guidance for focusing on outcomes or processes most proximal to desired outcomes.

In such exceptions, the following additional requirements could be considered.

- The performance measure is specified so that it requires providers to administer a specific PROM or clinical test at designated intervals and record the PROM or assessment value in the health record, not merely check off that it was administered.
- The developer submits a credible plan to implement the performance measure, collect data, and develop and test the outcome performance measure.

Recognizing the additional complexity of PRO-PMs, the TEP acknowledged that developing an outcome performance measure may not be immediately possible and that some flexibility to accept a performance measure focused on administering a PROM may be needed to begin collecting PROM data. However, an outcome measure is the goal; and a performance measure focused on the process of administering a PROM should only be considered in an exceptional circumstance. The proposed process measure should clearly specify that PROM values are collected not merely that it was administered; and there also should be a credible plan to develop the outcome measure.

Another issue is whether a performance measure focused on the process of administering a PROM could be considered for NQF endorsement. According to NQF criteria and prior guidance, outcomes and evidence-based processes proximal to outcomes are preferred for performance measurement. Typically, collecting data for assessment (e.g., lab value, vital signs, PROM) is quite distal to the outcome of interest. That is, there are multiple process steps between performing an assessment or collecting assessment data and the outcome – i.e., review the data; interpret the data correctly; identify appropriate treatment options; discuss data, treatment options, and recommendations with the patient; administer treatment according to protocol. Assessment is necessary but not sufficient to influence outcomes; and the evidence for the link between an assessment and outcome generally will only be indirect. Most evidence is typically focused on the link between the assessment value (e.g., BP >140/90) or treatments and the desired outcomes. However, NQF criteria do allow for an exception to the evidence criterion.

The primary purpose of a "process performance measure" focused on administrating a PROM is to facilitate use of the PROM to obtain the data needed to develop and test an outcome performance measure. Because developing, testing and endorsing such a process performance measure requires considerable resources, it should only be considered in an exceptional circumstance and where there is a credible plan to develop the outcome measure.

## Specification of the PRO-PM

**Recommendation 6.**

NQF should require measure specifications for PRO-PMs that include all the following: the specific PROM(s); standard methods, modes, and languages of administration; whether (and how) proxy responses are allowed; standard sampling procedures; the handling of missing data; and calculation of response rates to be reported with the performance measure results.

Performance measures used in accountability applications must be standardized. Therefore, developers must specify them in ways that will help to ensure consistent implementation across providers. Not unlike other performance measures, specifications should identify the data collection tool – i.e., the specific PROM(s) used to obtain the data for each patient (respondent). Specifications that are unique to PRO-PMs include standard methods, modes, and languages of administration, whether (and if so, how) proxy responses are allowed, standard sampling procedures, how missing data are handled; and how response rates are calculated and reported with the performance measure results.

## Reliability and Validity of Both the PROM and the PRO-PM

**Recommendations 7-8.**

**7.** NQF should require testing for PRO-PMs that demonstrates the reliability of both the underlying PROM in the target population and the performance measure score.

**8.** NQF should require testing for PRO-PMs that demonstrate the validity of both the underlying PROM in the target population and the performance measure score. Empirical validity testing of the performance measure is preferred. If empirical validity testing of the performance measure is not possible, a systematic assessment of face validity should be accomplished with experts other than those who created the measure, including patients reporting on the PROM, and this assessment should specifically address the approach to aggregating the individual PROM values.

As already noted, NQF endorses performance measures; it does not endorse instruments or scales (i.e., the PROM) alone. However, the PROM values are the data used in the performance measure, so the psychometric soundness of the PROMs specified for use in the performance measures is crucial to the reliability and validity of the PRO-PM. The Expert Panel agreed that reliability and validity of the PROM is necessary but not sufficient to ensure reliability and validity of the PRO-PM; therefore, it recommended that testing for both the PROM and the PRO-PM are needed. Approaches to reliability and validity testing, risk adjustment, and analyses of potential threats to validity were discussed in a commissioned paper on methods issues related to PRO-PMs (available here).

NQF criteria currently allow for testing reliability and validity for either the critical data elements used in the performance measure or for the computed performance measure score. In the case of the PRO-PM, a critical data element is the PROM value.

PROMs have traditionally been developed for group comparisons in research rather than for decisions about individual patients or service recipients. In a research context, investigators usually assign subjects randomly to treatment and control groups; by contrast, in healthcare settings and systems, patients are not randomly assigned to a provider of healthcare or support services. The primary question is whether demonstrated reliability and validity of the PROM is sufficient in itself to assume reliability and validity of the performance measure. NQF can consider two approaches to deal with this issue.

1. Accept reliability and validity of the PROM in the target population as meeting NQF criteria for reliability and validity testing at the data element level as long as the additional issues related to threats to validity are tested and analyzed for the performance measure score (i.e., exclusions, risk adjustment, discriminating performance comparability if multiple PROMs are used).

2. Require reliability and validity testing of the computed performance measure score *in addition to* providing evidence of reliability and validity of the PROM in the target population. The related threats to validity must also be addressed (i.e., exclusions, risk adjustment, discriminating performance comparability if multiple PROMs are used).

The primary advantage of the first approach is that measure developers can expend fewer resources for measure testing. The primary disadvantage of the first approach is less confidence in the results of the performance measure. The advantages and disadvantages of the second approach are the opposite.

The Expert Panel agreed that the second approach is more appropriate in the context of performance measures that NQF endorses for purposes of accountability and performance improvement. Further, the impact on resources for additional testing is not substantial, given the need to address threats to validity, such as differences in case mix or use of multiple PROMs, with either approach. For example, developers could use the data needed to develop and test risk adjustment to conduct reliability testing of the performance measure such as a signal-to-noise analysis. Therefore, a requirement for reliability testing of the performance measure would not present an additional burden on developers.

Validity testing of the performance measure score would require additional data to test hypothesized relationships such as correlation with another performance measure or comparison of performance scores for groups known to differ on quality. NQF criteria currently allow a systematic assessment of face validity of the performance measure score as an indicator of quality. Because developers can specify the performance measure to aggregate individual PROM values in various ways, the validity of results for indicating quality could differ as well. Ideally, developers would conduct empirical validity testing. If that is not possible, then they should evaluate face validity systematically with experts, including patients reporting on the PROM, other than those who created the measure.

## Missing Data and Response Rates

**Recommendations 9.**

NQF should require analysis of missing data and response rates to demonstrate that potential problems in these areas do not bias the performance measure results.

Missing data is an important consideration when using PROM data for performance measurement. This issue encompasses missing responses on a multi-item scale; missing responses from eligible patients and its impact on potential response bias; missing information because of exclusions; and using proxies to mitigate potential missing responses. Systematic missing data affects validity. Processes must be in place to safeguard against these exclusions and biases, and more robust engagement strategies are needed over time to prevent or mitigate poor response rates. NQF criteria for validity currently address exclusions, and missing data is often an explicit or implicit exclusion. Because missing data are likely to be more prevalent with PRO-PMs than with performance measures based on clinical data, developers should address this problem explicitly in measure specifications and testing the PRO-PM, which will be evaluated by NQF.

## Feasibility

**Recommendation 10.**

NQF’s feasibility criterion should consider the burden to both individuals providing PROM data (patients, service recipients, respondents) and the providers whose performance is being measured. The electronic capture criterion needs to be modified to include PROM data, not just clinical data.

The general principles of feasibility for a performance measure apply to PRO-PMs. Burden of data collection usually applies to the healthcare or service provider whose performance is being measured; however, the unique issue that needs to be considered with PRO-PMs is the potential burden to the individuals who are providing the PROM data. Burdens to both individuals and the providers delivering healthcare or support services will influence response rates, missing data, and ultimately the reliability and validity of a performance measure. Flexibility to decrease burden, such as collecting PROM data through tools developed in multiple languages and applying different methods and modes of administration, is desirable.

As with all performance measures, data collection and reporting for PRO-PMs may present a variety of costs to the providers whose performance is being measured. Such costs may involve expenditures on infrastructure such as computers and programming; they may, in some cases, entail paying licensing or other fees for proprietary instruments or measures. A potential difference between PRO-PMs and other performance measures regarding infrastructure is that, currently, PROMs are not widely in use and the needed information technology infrastructure is less advanced than that of electronic health records.

When considering burdens, developers and NQF need to weigh them against benefits. Obtaining PROM data is not merely a process to collect data for performance measurement. Rather, providers can use the PROM to assess patient status or response to intervention, plan and manage care or services, provide feedback for self-management, and engage patients in SDM (as desired by patients). The benefits of performance measurement and reporting are widely accepted. As with other performance measures, the burden of data collection does not stop performance measurement; rather, it should serve as an impetus to find more efficient ways to collect PROM data and to use resources for performance measurement on PRO-PMs that meet NQF criteria.

### Usability and Use

As with any NQF-endorsed measure, an NQF-endorsed PRO-PM is intended for use in both accountability and improvement applications. The primary indications of whether a performance measure can be applied for these purposes are whether it is in use and whether it is making a difference. At the time of initial NQF endorsement, of course, usability may be only theoretical. The performance measure may have a rationale and plans for use in accountability and improvement activities. On subsequent review for endorsement maintenance, however, NQF requires information on use and data on improvement. NQF also requests public comment on experiences with using the performance measure.

## Next Steps and Future Directions

This project provided a forum for dialogue among numerous and diverse stakeholders to address difficult conceptual, methodological, and practical issues. The aim was to hasten the endorsement and ultimately the implementation of PRO-based performance measures for use in accountability programs and performance improvement initiatives. The guiding principles articulated above and the detailed pathway (Figure 2) of taking a PRO to a PRO-PM are intended to steer work in the field in ways that help to ensure a more person-centered approach. This report begins to lay a roadmap to get the nation there.

Recommendations regarding NQF endorsement criteria require incorporation into NQF criteria, guidance documents, submission form, and processes for evaluation. Some of the recommendations (3, 4, and 5) require further examination in terms of implications for evaluation and endorsement of all performance measures.

PROMs that are at a state of readiness to address performance measure gap areas most meaningful to patients, such as functional status, could be taken down the recommended pathway to develop a PRO-PM and then through the NQF endorsement process. NQF anticipates incorporating PRO-PMs across the domains identified in this report into the broader measure endorsement agenda. PRO-PMs can be submitted for relevant condition-specific topic areas such as cardiovascular or pulmonary, as well as crosscutting areas such as functional status or care coordination.

### Issues for Further Work

Nevertheless, some pressing methods issues require further examination. The examples given here are high-priority needs to fill. First, identifying and evaluating best practices for using proxy respondents are important next steps; the goal is not to exclude from our assessments various disadvantaged populations, such as frail elders or children, who may be unable to respond to PROMs on their own.

Second, PROs may be evaluated through different PROMs (instruments); demonstrating the equivalency of the data from different PROMs warrants careful attention. Of particular concern is the trade-off between allowing implementers as much flexibility as possible without sacrificing validity and enhancing the ability of users to do meaningful comparisons. Third, viable solutions are needed to overcome barriers to calibrating multiple individual-level PROMs (i.e., "disparate" data sources) to a standard scale. Finally, some considerations will arise as use of PROMs and PRO-PMs expands and evolves. These include the advisability and utility of calculating composite endpoints or combining PRO-PMs salient to a particular domain such as health-related quality of life or health behaviors. Having such a broad picture of the outcomes reflected in the PRO-PMs strongly appeals to consumers who want a complete picture of health and well-being.

Using information technology to enable the widespread collection and use of PRO-based performance measures requires further exploration to capitalize fully on existing and future infrastructure. Technology can increase response rates by allowing individuals or their proxy respondents to provide responses to PROMs from home or elsewhere via telephone, computer tablet, or web-based systems. Technology permits scanning paper and pencil responses; this also allows for real-time scoring and giving feedback to respondents. Computers are an essential technology for real-time application of item response theory in computer adaptive testing, which allows more efficient administration of PROMs and calibration of multiple instruments to a standard scale.

Integrating PROMs into electronic health records (EHRs) can facilitate their use for patient-centered care management and also provide data for performance improvement, but implementers must take account of several social and technical factors. Some PROMs, such as those focused on people's experience with care, may not be appropriate to include in EHRs because current tools and approaches are based on the premise of anonymity. Incorporating data provided by patients into the health record may increase their sense of ownership of the record; doing so may also raise demands for extracting information and for providing data.

Data standards are needed before PROM data can be fully incorporated into EHRs. Formulating such standards requires making decisions about aspects of capturing PROM data such as the following: source of the information (e.g., self or proxy); specific PROM instrument; method and mode of data collection; PROM value or response; and dates on which information was captured and scores were computed. In addition, how PROM data might be used in clinical practice needs to be clearly specified, including how best to display results and when and how alerts should appear. This is an opportune time to include PROMs in EHRs and leverage the resources being directed to adoption of EHRs through the Medicare EHR Incentive Program referred to as "Meaningful Use."

In closing, the path forward toward NQF endorsement of PRO-based performance measures (PRO-PMs) is promising. This project has built on many years of exemplary work in the field of patient-reported outcomes. It now lays out concrete steps to move measurement and use of such data to the forefront of accountability and performance improvement.

## References

1. Institute of Medicine, *To Err Is Human: Building a Safer Health System*, Washington, DC: National Academy Press; 1999.
2. Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the 21st Century*, Washington, DC: National Academy Press; 2001.
3. Fisher ES, Wennberg DE, Stukel TA, et al., The Implications of Regional Variations in Medicare Spending (Parts 1 & 2), *Ann Intern Med*, 2003;138(4):273-298.
4. McGlynn EA, Asch SM, Adams J, et al., The Quality of Health Care Delivered to Adults in the United States, *NEJM*, 2003;348(26):2635-2645.
5. The Commonwealth Fund Commission on a High Performance Health System, *Why Not the Best? Results from the National Scorecard on U.S. Health System Performance, 2011*, New York: The Commonwealth Fund; 2011.
6. Agency for Healthcare Research and Quality (AHRQ), *National Healthcare Disparities and Quality Reports*, Rockville, MD: AHRQ; 2010. Available at http://www.ahrq.gov/qual/qrdr10.htm. Last accessed October 2012.
7. Patient Protection and Affordable Care Act, PL No. 111-148, Sec. 3011 (2010).
8. Department of Health and Human Services (HHS), *Report to Congress: National Strategy for Quality Improvement in Health Care*, Washington, DC: HHS; 2011. Available at http://www.healthcare.gov/law/resources/reports/nationalqualitystrategy032011.pdf. Last accessed October 2012.
9. Department of Health and Human Services (HHS), *Attachment to the Annual Report to Congress: National Strategy for Quality Improvement in Health Care: Agency-Specific Quality Strategic Plans*, Washington, DC: HHS; 2012. http://www.ahrq.gov/workingforquality/nqs/nqsplans.pdf. Last accessed October 2012.
10. Remmers C, Hibbard J, Mosen DM, et al., Is Patient Activation Associated with Future Health Outcomes and Healthcare Utilization among Patients with Diabetes? *J Ambul Care Manage*, 2009;32(4):320-7.
11. Weinstein JN, Clay K, Morgan TS, Informed Patient Choice: Patient-Centered Valuing of Surgical Risks and Benefits, *Health Affairs*, 2007;26(30):726-730.
12. Agency for Healthcare Research and Quality (AHRQ), *Guide to Patient and Family Engagement Environmental Scan Report*, Rockville, MD: AHRQ; 2012. Available at http://www.ahrq.gov/qual/ptfamilyscan/ptfamilyscan.pdf. Last accessed October 2012.
13. Conway J, Johnson B, Edgman-Levitan S, et al., *Partnering with Patients and Families to Design a Patient- and Family-Centered Health Care System: A Roadmap for the Future: A Work in Progress*, Bethesda, MD: Institute for Family-Centered Care; 2006. Available at http://www.ipfcc.org/pdf/Roadmap.pdf. Last accessed October 2012.
14. National Technology Transfer and Advancement Act of 1995, PL 104-113, 15 USC 3701 (1996).
15. U.S. Food and Drug Administration, Guidance for Industry, Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims, *Fed Regist.* 2009;74(35):65132-133. Available at http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Last accessed October 2012.
16. Cella D, Hahn EA, Jensen SE, et al. *Methodological Issues in the Selection, Administration and Use of Patient-Reported Outcomes in Performance Measurement in Health Care Settings.* Prepared by Northwestern University. Washington, DC:NQF; 2012. Available at http://www.qualityforum.org/Projects/n-r/Patient-Reported_Outcomes/Patient-Reported_Outcomes.aspx#t=2&s=&p=2%7C. Last accessed January 2013.

17. Deutsch A, Smith L, Gage B, et al. *Patient-Reported Outcomes in Performance Measurement.* Prepared by RTI International and Brookings Institution. Washington, DC:NQF; 2012. Available at http://www.qualityforum.org/Projects/n-r/Patient-Reported_Outcomes/Patient-Reported_Outcomes.aspx#t=2&s=&p=2%7C. Last accessed January 2013.
18. Informed Medical Decisions Foundation, *What is Shared Decision Making?* Available at http://informedmedicaldecisions.org/what-is-shared-decision-making/. Last accessed October 2012.
19. Elwyn G ,O'Connor A ,Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process, *BMJ,* 2006;333:417.
20. Sepuch KR, Stacey D, Clay CF, et al. Decision quality instrument for treatment of hip and knee osteoarthritis: a psychometric evaluation, *BMC Musculoskelet Disord,* 2011;12:149.

## Appendix A—Glossary

**Health behavior:** Behaviors expressed by individuals to protect, maintain or promote their health status. For example, proper diet, and appropriate exercise are activities perceived to influence health status.[1]

**Health-related quality of life (HRQoL):** A multi-dimensional concept that includes domains related to physical, mental, emotional and social functioning. It goes beyond direct measures of population health, life expectancy and causes of death, and focuses on the impact health status has on quality of life.[2]

**Patient-reported outcome (PRO)**: The concept of any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else. PRO domains included in this project encompass:

- health-related quality of life (including functional status);
- symptom and symptom burden;
- experience with care; and
- health behaviors.

**Performance measure**: Numeric quantification of healthcare quality for a designated accountable healthcare entity, such as hospital, health plan, nursing home, clinician, etc.

**PRO measure (PROM)**: Instrument, scale, or single-item measure used to assess the PRO concept as perceived by the patient, obtained by directly asking the patient to self-report (e.g., PHQ-9).

**PRO-based performance measure (PRO-PM)**: A performance measure that is based on PROM data aggregated for an accountable healthcare entity (e.g., percentage of patients in an accountable care organization whose depression score as measured by the PHQ-9 improved).

**Quality of Life:** An individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the person's physical health, psychological state, level of independence, social relationships, and their relationships to salient features of their environment.[3]

---

[1] U.S. National Library of Medicine. MESH Term: Health Behaviors. Available at http://www.ncbi.nlm.nih.gov/mesh?term=health%20behaviors. Last accessed January 2013.

[2] U.S. Department of Health and Human Services. Health-Related Quality of Life and Well-Being. Available at http://www.healthypeople.gov/2020/about/qolwbabout.aspx. Last accessed January 2013.

[3] Bowling A. "Health-Related Quality of Life: A Discussion of the Concept, its Use and Measurement Background: the Quality of Life." Presentation to the Adapting to Change Core Course, Washington, DC; September 1999. Available at http://info.worldbank.org/etools/docs/library/48475/m2s5bowling.pdf. Last accessed January 2013.

## Appendix B—Expert Panel Roster

***Patricia Brennan, RN, PhD, FAAN, FACMI (Co-Chair)**
University of Wisconsin-Madison, Madison, WI

***Joyce Dubow, MUP (Co-Chair)**
AARP, Washington, DC

**Richard Bankowitz, MBA, MD, FACP**
Premier healthcare alliance, Washington, DC

**Ethan Basch, MD, MSc**
University of North Carolina at Chapel Hill, Chapel Hill, NC

**Jim Bellows, PhD, MPH**
Kaiser Permanente, Oakland, CA

**Laurie Burke, RN**
Food and Drug Administration, Silver Spring, MD

**Jennifer Eames-Huff**
Pacific Business Group on Health, San Francisco, CA

***Stephan Fihn, MD, MPH**
Veterans Health Administration, Seattle, Washington

**Floyd Fowler, PhD**
Informed Medical Decision Making Foundation, Boston, MA

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

**Theodore Ganiats, MD**
University of California San Diego Health System, La Jolla, CA

***Kate Goodrich, MD**
Centers for Medicare & Medicaid Services, Washington, DC

**Judith Hibbard, DrPH**
University of Oregon, Eugene, OR

**Dennis Kaldenberg, PhD**
Press Ganey Associates, South Bend, IN

**Irene Katzan, MD, MS**
Cleveland Clinic, Cleveland, OH

***Lewis Kazis, ScD**
Boston University School of Public Health, Boston, MA

**Uma Kotagal, MSc**
Cincinnati Children's Hospital Medical Center, Cincinnati, OH

**Kevin Larsen, MD, FACP**
Office of the National Coordinator for Health Information Technology, Washington, DC

**Kathleen Lohr, PhD, MPhil, MA**
RTI International, Chapel Hill, NC

**Elizabeth Mort, MD**
Massachusetts General Hospital, Boston, MA

**Charles Moseley, EdD**
National Association of State Directors of Developmental Disabilities Services, Alexandria, VA

***Eugene Nelson, DSc, MPH**
Dartmouth- Hitchcock Medical Center, Lebanon, NH

**Kenneth Ottenbacher, PhD, OTR**
The University of Texas Medical Branch at Galveston, Galveston, TX

***Greg Pawlson, MD, MPH**
BlueCross BlueShield Association, Washington, DC

**Eleanor Perfetto, PhD**
Pfizer, Washington, DC

**Collette Pitzen, BSN, RN, CPHQ**
Minnesota Community Measurement, Minneapolis, MN

**Cheryl Powell**
Centers for Medicaid & Medicare Services, Baltimore, MD

**David Radley, PhD**
Institute for Healthcare Improvement, Cambridge, MA

**Ted Rooney, RN, MPH**
Maine Quality Counts, Manchester, ME

**Debra Saliba, MD, MPH**
UCLA Borun Center, Los Angeles VA Medical Center, The RAND Corporation, Los Angeles, CA

**Marcel Salive, MD, MPH**
National Institutes of Health, Rockville, MD

**Barbara Summers, PhD, RN, FAAN**
University of Texas-MD Anderson Cancer Center, Houston, TX

**Kalahn Taylor-Clark, PhD, MPH, BA**
National Partnership for Women & Families, Washington, DC

**Mary Tinetti, MD**
Yale New Haven Health System, New Haven, CT

**Phyllis Torda, MA**
National Committee for Quality Assurance, Washington, DC

**John Wasson, MD**
Dartmouth Medical School, Lebanon, NH

**Robert Weech-Maldonado, PhD**
University of Alabama at Birmingham, Birmingham, AL

**Linda Wilkinson, MBA**
Dartmouth Hitchcock Medical Center, Lebanon, NH

**Albert Wu, MD, MPH**
Johns Hopkins Health System, Baltimore, MD

“ * ” Indicates a member of the Planning Committee

## Project Staff

**Helen Burstin, MD, MPH**
Senior Vice President, Performance Measures

**Karen Adams, PhD, MT**
Vice President, National Priorities

**Heidi Bossley, MSN, MBA**
Vice President, Performance Measures

**Karen Pace, PhD, MSN**
Senior Director, Performance Measures

**Eugene Cunningham, MS**
Project Manager, Performance Measures

**Evan M. Williamson, MPH, MS**
Project Analyst, Performance Measures

## Appendix C—Characteristics for Selecting PROMs

### Table 4[4]. Important characteristics and best practices to evaluate and select PROs for use in performance measures[279,284]

| **Characteristic** | **Specific issues to address for performance measures** | **Example: The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)[354] for use in hip arthroplasty** |
|---|---|---|
| **1. Conceptual and Measurement Model** | | |
| A PRO measure should have documentation defining and describing the concept(s) included and the intended population(s) for use. | • Target PRO concept should be a high priority for the healthcare system and patients. Patient engagement should define what is an important concept to patients.<br>• Target PRO concept must be actionable in response to the healthcare intervention. | • Factorial validity of the physical function and pain subscales has been inadequate.[355] |
| There should be documentation of how the concept(s) are organized into a measurement model, including evidence for the dimensionality of the measure, how items relate to each measured concept, and the relationship among concepts. | | |
| **2.Reliability** | | |
| The degree to which an instrument is free from random error. | | |
| **2a.** ***Internal consistency*** *(multi-item scales)* | Classical Test Theory (CTT):<br>▪ reliability estimate ≥ 0.70 for group-level purposes<br>▪ reliability estimate ≥ 0.90 for individual-level purposes<br>Item Response Theory:<br>• item information curves that demonstrate precision [181]<br>• a formula can be applied to estimate CTT reliability | • Cronbach alphas for the three subscales range from 0.86 to 0.98.[356-358] |
| **2b.** ***Reproducibility*** *(stability over time)*<br>▪ type of test-retest estimate depends on the response scale (dichotomous, nominal ordinal, interval, ratio) | | • Test-retest reliability has been adequate for the pain and physical function subscales, but less adequate for the stiffness subscale.[358] |

---

[4] This table is adapted from recommendations contained within a report from the Scientific Advisory Committee of the Medical Outcomes Trust and a report submitted to the PCORI Methodology Committee. The recommendations from these sources have been adapted to enhance relevance to PRO selection for performance measurement.

| Characteristic | Specific issues to address for performance measures | Example: The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)[354] for use in hip arthroplasty |
|---|---|---|
| **3. Validity** | | |
| The degree to which the instrument reflects what it is supposed to measure. | • There are a limited number of PRO instruments that have been validated for performance measurement.<br>• PRO instruments should include questions that are patient-centered. | |
| ***3a. Content Validity*** | | |
| The extent to which a measure samples a representative range of the content. | | |
| A PRO measure should have evidence supporting its content validity, including evidence that patients and/or experts consider the content of the PRO measure relevant and comprehensive for the concept, population, and aim of the measurement application. | | • Development involved expert clinician input, and survey input from patients,[359] as well as a review of existing measures. |
| Documentation of qualitative and/or quantitative methods used to solicit and confirm attributes (i.e., concepts measured by the items) of the PRO relevant to the measurement application. | | |
| Documentation of the characteristics of participants included in the evaluation (e.g., race/ethnicity, culture, age, socio-economic status, literacy). | | |
| Documentation of sources from which items were derived, modified, and prioritized during the PRO measure development process. | | |
| Justification for the recall period for the measurement application. | | |
| ***3b. Construct and Criterion-related Validity*** | | |
| A PRO measure should have evidence supporting its construct validity, including:<br>• documentation of empirical findings that support predefined hypotheses on the expected associations among measures similar or dissimilar to the measured PRO<br>• documentation of empirical findings that support predefined hypotheses of the expected differences in scores between "known" groups | | • Patient ratings of satisfaction with arthroplasty were correlated with WOMAC scores in the expected direction.[22,360,361] |
| A PRO measure should have evidence that shows the extent to which scores of the instrument are related to a criterion measure. | | |

| Characteristic | Specific issues to address for performance measures | Example: The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)[354] for use in hip arthroplasty |
|---|---|---|
| ***3c. Responsiveness*** | | |
| A PRO measure for use in longitudinal initiatives should have evidence of responsiveness, including empirical evidence of changes in scores consistent with predefined hypotheses regarding changes in the target population. | • If a PRO measure has cross-sectional data that provides sufficient evidence in regard to the reliability (internal consistency), content validity, and construct validity but has no data yet on responsiveness over time (i.e., ability of a PRO measure to detect changes in the construct being measured over time), would you accept use of the PRO measure to provide valid data over time in a longitudinal study if no other PRO measure was available? | • Demonstrates adequate responsiveness and ability to detect change in response to clinical intervention.[362] |
| | • Important to emphasize responsiveness because there is an expectation of consequences. Need to be able to demonstrate responsiveness if action is to be taken. | |
| | • PRO must be sensitive to detect change in response to the specific healthcare intervention | |
| **4. Interpretability of Scores** | | |
| A PRO measure should have documentation to support interpretation of scores, including:<br>• what low and high scores represent for the measured concept<br>• representative mean(s) and standard deviation(s) in the reference population<br>• guidance on the minimally important difference in scores between groups and/or over time that can be considered meaningful from the patient and/or clinical perspective | • If different PROs are used, it is important to establish a link or cross-walk between them.<br>• Because the criteria for assessing clinically important change in individuals does not directly translate to evaluating clinically important group differences,[327] a useful strategy is to calculate the proportion of patients who experience a clinically significant change[271,327] | • Availability of population-based, age- and gender-normative values[363]<br>• Availability of minimal clinically important improvement values[364]<br>• Can be translated into a utility score for use in economic and accountability evaluations[365] |

| **Characteristic** | **Specific issues to address for performance measures** | **Example: The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)[354] for use in hip arthroplasty** |
|---|---|---|
| **5.Burden** | | |
| The time, effort, and other demands on the respondent and the administrator. | • In a busy clinic setting, PRO assessment should be as brief as possible, and reporting should be done in real-time.<br>• Patient engagement should inform what constitutes "burden." | • Short form available[366]<br>• Average time to complete mobile phone WOMAC = 4.8 minutes[367] |
| **6. Alternatives modes and methods of administration** | • The use of multiple modes and methods can be useful for diverse populations. However, there should be evidence regarding their equivalence. | • Validated mobile phone and touchscreen based platforms[368,369] |
| **7. Cultural and language adaptations** | • The mode, method and question wording must yield equivalent estimates of PRO measures. | • Available in over 65 languages[370] |
| **8. Electronic health records (EHR)** | Critical features:<br>▪ interoperability<br>▪ automated, real-time measurement and reporting<br>▪ sophisticated analytic capacities | ▪ Electronic data capture may allow for integration within EHR[367] |

National Quality Forum
1030 15th St NW, Suite 800
Washington, DC 20005
http://www.qualityforum.org

Contract HHSM-500-2009-00010C
Task order 14.2

ISBN 978-1-933875-44-6

VIEWPOINT

# Standards for Patient-Reported Outcome–Based Performance Measures

**Ethan Basch, MD, MSc**
School of Medicine and School of Public Health, University of North Carolina, Chapel Hill.

**Phyllis Torda, MA**
National Committee for Quality Assurance, Washington, DC.

**Karen Adams, PhD**
National Quality Forum, Washington, DC.

**Viewpoint** page 137

**Supplemental content** at jama.com

**The science of collecting** and analyzing patient-reported outcomes (PROs) is well developed in clinical research.[1] Almost a quarter of US drug labels include information about treatment benefits based on patient symptom questionnaires.[2] International organizations and regulatory agencies have developed methodological standards for the use of PROs in clinical trials.[1,3]

But this work has largely occurred outside the workflows of those delivering care and the vocabulary of performance measurement. Although data about patients' impressions of or experiences with care delivery (ie, satisfaction) are routinely collected, reports about symptoms, functional status, or quality of life are not.

To close this gap, recent initiatives by several major US organizations involved with the development, endorsement, and implementation of performance measures have converged on approaches for collecting, analyzing, and reporting outcomes that patients notice and care about (ie, patient-centered).

The National Quality Forum (NQF), the National Committee for Quality Assurance (NCQA), and a number of US medical specialty societies have moved forward with initiatives to promote the use of PROs as a basis for performance measurement. These efforts recognize the need to move from a PRO measure (PROM) to a PRO-based performance measure (PRO-PM), which specifies how patient-reported data are aggregated and interpreted to reflect performance (**Box**).

In January 2013, the NQF produced a report outlining a pathway to the endorsement of PRO-PMs and criteria for evaluating them for endorsement as US national voluntary consensus standards.[4] This report was based on a series of multidisciplinary stakeholder workshops that included patient advocates, clinicians, experts in performance assessment, and PRO experts and was informed by 2 commissioned methods papers and a public comment period.

Approaches used in several existing performance evaluation programs that integrate PROs also informed the report. Some examples include a program for universal reporting of symptoms and functional status by patients in England after selected elective surgeries[5]; administration of a symptom and functional status questionnaire to patients enrolled in Medicare Advantage plans[6]; and a state-sponsored program for collection of patient-reported depression scores by primary care and psychiatric practices across Minnesota.[7]

### Box. Definitions Related to Assessing Performance Using Patient-Reported Outcome Measures

**Patient-reported outcome (PRO):** The concept of any report of the status of a patient's health condition that comes directly from the patient (or in some cases a caregiver or surrogate), without interpretation of the patient's response by a clinician or anyone else. An example is the concept of depression.

**PRO measure (PROM):** An instrument, scale, or single-item measure used to assess the PRO concept as perceived by the patient, obtained by directly asking the patient (or in some cases a caregiver or surrogate) to self-report. An example is the Patient Health Questionnaire-9 (PHQ-9).

**PRO-based performance measure (PRO-PM):** A performance measure that is based on PROM data aggregated for an accountable health care entity. An example is the proportion of patients with depression or dysthymia and an initial PHQ-9 score >9 who after 6 months of management from mental health professionals have a PHQ 9 score <5 at follow-up.

A pathway has been generated through this effort for developing methodologically robust and actionable PRO-PMs (eFigure in Supplement). This pathway begins with establishing a rationale for using a PRO to measure performance in a particular context—including a process for identifying outcomes that are meaningful to patients. It carries through to selecting or developing an appropriate PROM, then to identifying and testing an approach for aggregating and interpreting the patient-reported data as a PRO-PM. The PRO-PM can then be used for purposes of accountability and quality improvement. In addition, there is periodic evaluation of the PRO-PM as part of ongoing feedback and refinement.

These recommendations are already being used. The NCQA has prior experience with PRO performance measurement through its involvement with the Medicare Advantage program (and provided input in the development of the NQF recommendations). Now, the NCQA is working with the US Office of the National Coordinator for Health Information Technology and the Centers for Medicare & Medicaid Services (CMS) to identify PRO measurement strategies supported by electronic health records. This includes developing consensus on which PROMs are appropriate in specific clinical situations; the timing of administration of PROMs; definitions of meaningful improvements; and risk adjustment methods.

A number of professional organizations are also developing condition-specific PRO-PMs. For example, the American Society of Clinical Oncology is creating PRO-PMs for specific symptoms and functional status relevant to the treatment of cancers toward implementation in its Quality Oncology Practice Initiative network, and is basing its work on the NQF model. The Society of Thoracic Surgeons and American College of Cardiology

**Corresponding Author:** Ethan Basch, MD, MSc, Cancer Outcomes Research Program, School of Medicine and School of Public Health, University of North Carolina, 170 Manning Dr, Chapel Hill, NC 27514 (ebasch@med.unc.edu).

recently developed a registry including assessment of patient-reported health-related quality of life as part of a CMS National Coverage Determination program for percutaneous transcatheter aortic valve replacement.

The American Medical Association–convened Physician Consortium for Performance Improvement (PCPI) recently hosted a technical expert panel to identify best practices for developing PRO-PMs. Similar to the NQF recommendations, key PCPI themes included the importance of engaging patients and consumers to help identify quality performance problems and corresponding meaningful outcomes that are amenable to change by health care intervention; to determine when PROs are appropriate; and to test the clarity, interpretability, and psychometric properties of potential PRO-PMs. Pilot testing was viewed as essential to evaluate feasibility of a measurement strategy and to obtain data for developing and testing the PRO-PM. The panel harmonized its work with the NQF recommendations through a collaborative effort.

Challenges for early PRO-PM programs will include the cost of new infrastructure and pilot testing; the logistics of collecting data from patients in a "real world" setting while minimizing missing data; and the development of analytic techniques that risk adjust and yield results that are meaningful to clinical practice. As programs move forward, these challenges should be periodically reconsidered.

A rationale for PRO-PM programs is that better symptom control and quality of life are associated with reduced costs and use of medical services and improved medication compliance, patient satisfaction, and survival.[8,9] Availability of PRO-PM information will help clinicians evaluate whether they are optimizing symptom control and learn how they might improve; can allow health systems or quality assessment programs to identify practices for which educational interventions or outreach programs could be beneficial; and, ultimately, may assist patients in understanding which practices best manage issues related to symptoms and functioning.

Now that a pathway has been outlined (with an understanding that it will be refined as experience accumulates from continuous feedback), the next step is wider integration of PROs into delivery of care. To date, this takes place in very few practices. In the near future, electronic health records will need to include fields to collect structured data and define data elements that can be represented with appropriate codes for this information.

Research funding and engagement of stakeholders across the quality enterprise—including payers, health systems, professional societies, researchers, and patient groups—are essential for fostering priority setting, rigorous measure development, and integration of PRO-PMs into accountability programs. Such efforts will help bring patients' perspectives to the center of care delivery and the center of performance measurement, where they belong.

## ARTICLE INFORMATION

**Conflict of Interest Disclosures:** All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

**Additional Contributions:** We thank Kathleen Blake, MD, MPH, for reviewing and providing feedback on this Viewpoint. Dr Blake did not receive compensation for her contribution.

## REFERENCES

**1**. Brundage M, Blazeby J, Revicki D, et al. Patient-reported outcomes in randomized clinical trials: development of ISOQOL reporting standards [published online September 2012]. *Qual Life Res*. 2012. doi:10.1007/s11136-012-0252-1.

**2**. Gnanasakthy A, Mordin M, Clark M, DeMuro C, Fehnel S, Copley-Merriman C. A review of patient-reported outcome labels in the United States: 2006 to 2010. *Value Health*. 2012;15(3):437-442.

**3**. US Department of Health and Human Services, Food and Drug Administration. Guidance for industry: patient-reported outcomes measures: use in medical product development to support labeling claims. www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Published December 2009. Accessed June 14, 2013.

**4**. National Quality Forum. Patient-reported outcomes (PROs) in performance measurement. www.qualityforum.org/Projects/n-r/Patient-Reported_Outcomes/Patient-Reported_Outcomes.aspx. Published January 10, 2013. Accessed June 14, 2013.

**5**. Devlin NJ, Appleby J. Getting the most out of PROMS. https://www.kingsfund.org.uk/sites/files/kf/Getting-the-most-out-of-PROMs-Nancy-Devlin-John-Appleby-Kings-Fund-March-2010.pdf. Published March 2010. Accessed June 14, 2013.

**6**. Jones N III, Jones SL, Miller NA. The Medicare Health Outcomes Survey program: overview, context, and near-term prospects. *Health Qual Life Outcomes*. 2004;2:33-43.

**7**. Minnesota Community Measurement. Minnesota health scores overview. www.mnhealthscores.org. Accessed June 14, 2013.

**8**. Harrison PL, Pope JE, Coberley CR, Rula EY. Evaluation of the relationship between individual well-being and future health care utilization and cost. *Popul Health Manag*. 2012;15(6):325-330.

**9**. Gotay CC, Kawamoto CT, Bottomley A, Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol*. 2008;26(8):1355-1363.

# Consumer-Led Evaluation Teams:
## A Peer-Led Approach to Assessing Consumer Experiences with Mental Health Services

Prepared for the National Empowerment Center
June 2009

**Sara Plachta-Elliott, MA, PhD (cand.)**
**Jonathan Delman, MPH, JD**

***CQI***
**Consumer Quality Initiatives, Inc.**

**Bringing the people's voice to behavioral health research…**
**and from research to practice.**

# Table of Contents

## Background

Consumer Quality Initiatives was asked by the National Empowerment Center to conduct an exploratory examination of consumer-led evaluation teams in order to identify best practices in consumer-run evaluation. Below we discuss our methodology for conducting this examination, the findings and our identified best practices for approaching this work.

### Consumer Quality Initiatives

Consumer Quality Initiatives, Inc. (CQI) is a mental health consumer-directed (51% of the board is consumers) and staffed non-profit research and evaluation organization whose mission is to develop opportunities for the meaningful involvement of consumers and family members in all aspects of mental health research and program evaluation (www.cqi-mass.org). By doing so, we aim to study issues that are relevant to the community, initiate changes to improve the system for all, and narrow the gap between research/evaluation and practice. http://www.cqi-mass.org/index.php4

In existence since January 1999, utilizes a Community-based Participatory Action Research (CBPR) framework, with an emphasis on protocols that are designed to impact policy and practice directly. The goal of CBPR is to identify a research topic of importance to the community (consumers, families) with the aim of combining knowledge and action for social change that improve community health. CQI engages in both research and evaluation, and has lead several NIMH grants to develop a methodology for conducting CBPR with the mental health community.

### National Empowerment Center

The mission of the National Empowerment Center Inc. is “to carry a message of recovery, empowerment, hope and healing to people who have been labeled with mental illness.” (http://www.power2u.org/what.html)

The National Empowerment Center is engaged in:
- Information and Referral
- Networking
- Conference Planning
- Lectures, Workshops and Consultation
- Publishing and Media
- Policy Issues
- Representation on National Boards
- Research
- Development of Educational Resources
- Development of Self-Help Resources

# Methodology

The primary selection criteria for indentifying “consumer-led evaluation teams” were that they:

1) be operated independently by consumers and/or family members, and thus not administratively managed by the mental health authority or agency,
2) evaluate mental health programs by learning about the experiences of program clients and/or ex-clients

Through an extensive internet search for consumer-led evaluation teams we identified several different entities that appeared to fall into this category, many of them in Pennsylvania because of the requirement that there be one in each county. These organizations’ websites were then examined to determine their suitability for inclusion in this analysis. Website content and online reports were reviewed to provide an initial understanding of each consumer-led evaluation team.

Of the twelve we identified, we contacted four organizations based on the variety of location, size, and approach they offered. We sent to the organization’s executive director (email and post) an introductory letter inviting him/her to participate in a one-hour phone interview, and three agreed to participate. In addition, we made the decision to profile our own organization in order to capture the kind of variety discussed above. After conducting the interview and writing a profile of their organizations, the Executive Directors were also given the opportunity to review a draft of the report to ensure accuracy in the information presented.

### Interview Guide and Interviews

In addition to information collected from the organizations' website, an interview guide was developed so that a comprehensive understanding of the organization, its history, and its work could be collected. The interview guide included questions on the following topics:

- *Organization*: Nonprofit/public status, history, mission, board/staffing, leadership, key partnerships.
- *Research and Evaluation*: Consumer satisfaction work, other research and evaluation, methodology (data collection, survey development, sample sizes), primary research and evaluation populations and levels of care, staff training, building internal research capacity.
- *Funding*: Past, current, future, funding mix.
- *Sharing Findings*: Sharing data with consumer and family member communities, providers, and policy makers; conferences, events, reports/papers, other.
- *Impact*: Programs, services, policy changes, noted improvements in quality of care, any formal attempts to measure impacts.

- *Future Directions*: Growing the organization, staffing, staff development, infrastructure, research and evaluation, impact.

During the telephone interviews, the interviewer (SPE) typed notes as close to verbatim as possible.

## Analysis

Once the interviews were complete, we reviewed and analyzed information collected from websites and interviews for key themes and to identify innovative practices. Once narratives about each organization were written, Executive Directors of the organizations were given the opportunity to review a draft before publication to ensure accuracy of the information contained in this report.

### Note on Terminology

For the purposes of this document, "consumer satisfaction" is a broad term covering a variety of ways in which consumers perceive the quality of services.

"Consumer satisfaction" covers a person's happiness/satisfaction with their services, and the kinds of questions asked tend to be very subjective. A good example of a question here is: "How satisfied are you with the staff here?"  You can provide a scale of responses (very satisfied, somewhat satisfied, somewhat dissatisfied, very dissatisfied) and/or ask for an open-ended response.

"Consumer satisfaction" also covers a person's perception of a specific aspect of their services. Good examples of questions are: "When you've told staff that you have a health concern, how often do they respond to that concern?" Again you can provide a scale of responses (never, sometimes, usually, always), and/or ask for an open-ended response. While still subjective, these questions have a more objective element since a person is asked to consider how often something has happened.  From a quality improvement perspective, these kinds of questions provide more "actionable" responses, meaning that someone who sees the data has a good idea of how respondents observe a specific issue, and the provider can take a specific action (eg, peers training staff, staff  “sensitivity" training, improved policies for staff for handling concerns).

## Overview of Consumer-Led Evaluation Teams

As noted by the President's New Freedom Commission on Mental Health report in 2003[1] and a growing body of literature on improving the quality of mental health care, a core flaw of our mental health system is the lack of consumer direction in program evaluation. Consumer directed must include defining outcomes, collecting and analyzing satisfaction data, and making quality improvement recommendations. As stated in a recent working paper, consumer involvement needs to:

> …go beyond satisfaction measures to formally involve consumers in system design, implementation, and monitoring. To be meaningful, the participation has to be sustained over time and focused on crucial elements of the program. Consumers will only participate if they feel their voice has an impact (Forquer & Sabin 2002).

The typical mental health consumer survey methodologies have not been effective in producing high quality data due to poor sample representations and lack of respondent openness. The consumer run personal interviewing apparatus, with its familiarity with the "community," is the critical component to developing a consumer-driven system where evaluation findings are effectively translated into practice.

In addition, consumer interviewers are able to gain greater information from other consumers about negative experiences they have had with services because "greater feelings of safety, trust, confidentiality, and privacy may have influenced the interviewees' disclosures to other clients… If clients with extremely negative experiences speak more freely with other clients, the result will be more valid feedback." (Clark et al. 1999, p.962-963). Consumers are in fact capable of taking a lead role in consumer satisfaction evaluation, but only if they are provided with the resources to support training and infrastructure. (Delman and Beinecke 2005)

Consumer evaluation is an excellent organizing tool. Consumer groups don't have systemic data on program quality, and without data it is difficult to effect change. Consumer data collection allows them to learn about the specifics of a program's performance and to work with administrators to set benchmarks for improvement. When consumers have data, they are in a better position to organize around an issue, and conversely to work with state administrators. Consumer evaluation also provides both training and jobs to consumers interested in evaluation. It's a stigma buster, demonstrating that consumers can handle "professional" jobs.

While many core components of consumer evaluation are shared, the replication of the consumer led evaluation team model has resulted in a variety of innovative approaches to consumers and family members collecting data from their peers about satisfaction with services and then reporting feedback to providers, funders, and government administrators in order to improve quality of care around satisfaction and consumer-direction.

---

[1] http://www.mentalhealthcommission.gov/

## Organizational Profiles

This report highlights the work of four consumer-led evaluation teams:

- Consumer Satisfaction Team, Inc. of Philadelphia;
- Vital Voices for Mental Health of Milwaukee.
- Consumer Quality Initiatives of Massachusetts;
- Consumer Quality Team of Maryland;

These four community-based organizations are run by consumers and family members, and focus exclusively on bringing the voices of consumers to mental health system planners and administrators. Although the organizations have many features in common, their differences highlight the variety of ways that consumer satisfaction work can be approached.

All but one of the organizations profiled is an independent 501(c)3 non-profit organization, and the other organization is currently housed under a mental health advocacy organization with plans to become an independent organization after a few years of operation.

### The Consumer Satisfaction Team, Inc. of Philadelphia

Mission: *"to ensure that specific publicly supported and funded services meet the expressed wishes and needs of the consumers of those services and that they promote maximal recovery of persons served."*

The Consumer Satisfaction Team, Inc. (CST) was created in 1990 in Philadelphia as the first consumer and family member run organization in the country to focus exclusively on quality assurance from the perspectives of mental health service consumers. CST has a large full-time staff of 27 that conducts consumer satisfaction interviews on a rotating basis at all facilities serving adults and children in the Philadelphia public mental health and substance abuse system. CST uses an open-ended interview method, then compiles notes from interviews and submits consumer feedback to governmental funding agencies and providers. CST staff also holds a biweekly accountability meeting with the funding agency administrators to determine whether each consumer's issue has been addressed. In addition to program site visits, consumers and family members can also call the CST to report issues.

### Vital Voices

Mission: *"to advocate for excellence in mental health services through the voice of the people being served."*

Vital Voices for Mental Health in Milwaukee County was conceived in the 1990s by a broad spectrum of stakeholders from the community, including consumers, family members, county agencies, and providers, as part of a master plan to improve mental health services in the county. The goal of the initiative was to get consumer voices into the system with the hope of improving services and satisfaction. Early on, the organizational founders decided that gathering consumer perspectives and reporting them to agency managers was not enough. A clear decision was made to fuse consumer satisfaction surveys with individual advocacy to ensure that consumers' individual needs were met by the system. Vital Voices spends about eight months of the year conducting individual consumer satisfaction interviews at a variety of providers and the remaining time administering the nationally recognized Mental Health Statistics Improvement Program (MHSIP) survey[2].

## Consumer Quality Initiatives

Mission: *"to develop opportunities for the meaningful involvement of consumers and family members in all aspects of mental health research and program evaluation. By doing so, we aim to study issues that are relevant to the community, initiate changes to improve the system for all, and narrow the gap between research/evaluation and practice."*

Consumer Quality Initiatives was established in 1998, originally as The Massachusetts Consumer Satisfaction Team, after the consumer empowerment and recovery movement marched on the state Medicaid office demanding greater consumer involvement in planning services, and specifically to fund a consumer-led evaluation team. The mental health managed care carve-out, Massachusetts Behavioral Health Partnership (MBHP), then contracted with CQI to evaluate services through personal interviews with clients, and that has remained CQI's largest contract.

CQI has three lines of work. First, there is evaluation and in particular consumer and family member satisfaction interviews and surveys to evaluate mental health and addiction services throughout Massachusetts. Second, CQI has taken the lead in several participatory action research projects, including several that are NIH funded, which allows it to learn about consumer perspectives outside of the arrangement of the current service system. Third, CQI consults with academia and state government on meaningfully involving consumers in research and evaluation.

CQI currently has 5 full-time staff, a few part-time staff, and several on-call consumer interviewers. While CQI's largest contract is with MBHP to do consumer satisfaction surveys at a variety of provider levels, CQI also holds contracts with the Massachusetts Department of Mental Health and sub-contracts with university research centers and private evaluation consulting firms.

## Consumer Quality Team of Maryland

Mission: *"…empowers individuals who receive services as partners with providers, policy makers and family members, to improve care in the public mental health system and ensure services meet the expressed needs of consumers."*

---

[2] http://www.mhsip.org/whatis.html

The newest of the consumer-led evaluation teams highlighted in this report, Consumer Quality Team of Maryland was funded in June 2006 by the Maryland Department of Mental Hygiene in response to a ten year campaign by consumers, providers and funding authorities led by the Mental Health Association of Maryland.

CQT interviewers make unannounced site visits to mental health facilities 3-6 times a year and conduct qualitative interviews with adult consumers to determine their satisfaction with services and to identify unmet needs or problems. CQT makes a verbal report to staff before leaving the site and follows up with a written report to the funding authority and the provider within ten days of the visit. Monthly meetings with representatives from each Core Service Agency, the Mental Hygiene Administration, and the provider associations allow CQT staff to go over reports in-depth and to make sure that individuals' complaints are being addressed.

Although the focus is on making sure individuals' needs are met, this process furnishes both real-time information about programs to the providers and funders, as well as information on emerging trends in people's needs, such as the loss of entitlements or social recreational programming as a result of recent budget cuts. The Funding Authority representative then gives CQT a written report stating what is being done to address concerns. The CQT staff has the opportunity to observe changes on subsequent site visits. Additionally, during every interview, consumers are given contact information for CQT and urged to call if difficulties continue. These calls are documented and followed in the same manner as the site visits.

## Findings

### A. Core Features of Consumer-Led Evaluation Teams

Interviews with directors of the four consumer-led evaluation teams and reviews of relevant materials revealed six core components shared by the organizations.

**Core Features of Consumer-Led Evaluation Teams**

1 – Created by consumers demanding input into service changes
2 – Consumer and family member run and staffed
3 – Strong connection with county/state agencies or managed care
4 – Collect consumer feedback through face-to-face peer-to-peer interviews
5 – Strong emphasis on provider and funder accountability
6 – Strong recovery perspective

## 1 Created by consumers demanding input into service changes

In each instance, the consumer-led evaluation team started in response to consumer demand for an increased consumer voice into service planning. State or county mental health authorities responded to these calls for involvement by allocating funds in their annual budgets for consumer satisfaction work. The closing of a state-run psychiatric hospital was the impetus for the Consumer Satisfaction Team, Inc. in Philadelphia to form. Consumers, family members, and advocates were calling for the City of Philadelphia to consider the preferences of former state hospital patients as they were moved to community living. The Philadelphia Department of Mental Health responded by allocating money for the creation of the Consumer Satisfaction Team.

During the1990s, the provision of mental health care in Massachusetts was moving to a Medicaid managed care model, and the consumer community soon realized that they were completely out of the mix in planning services. After a march on the Medicaid office and negotiations with the Medicaid office, the Department of Mental Health and the managed care company, an agreement was reached to create a consumer lead entity to evaluate Medicaid services from a consumer perspective, which lead to CQI.

In Milwaukee in the early 1990s, a representative group of consumers, family members, county officials and community agency representatives worked on a master plan for the county's mental health system. The people working on the plan went to Philadelphia to observe the work of the Consumer Satisfaction Team, Inc. to see if similar work could be done in Milwaukee. A plan was developed and the county issued a request for proposals (RFP). Two agencies submitted a joint proposal and were awarded the grant to develop the program, which eventually incorporated into a stand-alone organization to be run solely by people with mental illness and their family members.

Consumer Quality Team of Maryland was created in response to the efforts of a group of consumers, family members, providers, advocates, and state mental health agency administrators in the 1990s that pushed for the creation of a consumer-led evaluation team. Though it took years to secure funding, the partnership forged among these parties allowed the group to research existing programs and decide on the preferred method to collect consumer satisfaction information in Maryland in order to effect real-time change.

## 2 Consumer and family member run and staffed

These consumer-led evaluation teams are staffed by consumers and family members. The level of consumer directness of these programs is generally through the grant of authority to "do what needs to be done." However, the required and actual composition of the board varies, in terms of consumer and family member membership. Some boards of directors have a mandate requiring 51% of their boards be consumers. Other boards do not have such a mandate but strive to maintain a significant representation of consumers and family members.

### 3 Strong connection with county/state agencies or managed care

Though the impetus for the creation of consumer-led evaluation teams originated in the consumer advocacy movement, strategic partnerships with governmental agencies or private managed care organizations were essential for securing the funding and access to provider sites necessary to launch the organizations. Often, a few key individuals in government who truly believed in consumer empowerment partnered with consumer advocates to formulate plans and secure funding.

All of these consumer-led evaluation teams hold a contract with either the city, county or state mental health agency/authority to do surveys or interviews with consumers in the public mental health system. CST of Philadelphia's original contract is with the Office of Mental Health within the Department of Behavioral Health and Mental Retardation, but has expanded to include many other offices and departments. In Massachusetts, Consumer Quality Initiatives holds its largest contract with the managed care organization, but also contracts with the Massachusetts Department of Mental Health to do more focused needs assessments and case studies of the successful integration of empowered peer specialists in the services system.

Early on in their history and sometimes on an ongoing basis, these organizations faced resistance from providers regarding, among other things, the utility of the information that could be provided. Government's or managed care's support if not assertiveness in the face of these barriers are very important to the success of these teams. In the end, each organization must navigate their state or local system of mental health and addictions services to secure funding on an on-going basis.

### 4 Collect consumer feedback through face-to-face peer-to-peer interviews

All of the consumer-led evaluation teams conduct face-to-face interviews with consumers as their primary data collection method. The interviews usually take place at the program site where the consumer is receiving services. Vital Voices, CST and CQI also conduct surveys in people's homes, if the level of care received is community-based, such as case management. Some, such as CQI, have begun to use the phone in a limited way for people, such as parents of youth with mental health issues, who are too busy to meet in person.

### 5 Strong emphasis on provider and funder accountability

Ultimately, the purpose of collecting information from people using services about their satisfaction with and perspective of care is to ensure that consumers' needs are met. Consumer-led evaluation teams were created because most providers lacked the information they needed to be consumer-recovery oriented, and could in fact be completely insensitive to the perspectives of those seeking care and their preferences. The CST of Philadelphia has set a tone by having regular scheduled meetings with high levels of governmental authorities.

Moving the mental health system to a consumer-orientation requires providers and funders to be more accountable. The consumer-led evaluation teams highlighted in this report provide models for how consumers and family members can create an accountability process focused on bringing the experiences and needs of consumers to the attention of providers and policymakers.

### 6 Strong recovery perspective

Although not the direct focus of their work, all of the consumer-led evaluation teams commented on their commitment to a recovery perspective. Consumer-led evaluation teams aim to move the systems towards recovery-oriented services.

Having consumer interviewers out interviewing at program sites allows provider staff and program participants to see consumers in recovery who are helping to increase consumer voice. Providing consumers with meaningful work is an inherent, though not always explicit, goal of consumer-led evaluation teams.

In fact, more and more teams are using recovery oriented quantitative measures. For example, the Commonwealth of Pennsylvania funded the Consumer Satisfaction Team Alliance of Pennsylvania the use of the Recovery-Oriented Systems Indicators[3] (ROSI) statewide, both by mail and through personal interviewers of a smaller groups of individuals. In addition, CQI has incorporated aspects of the RSA[4] (Recovery Self-Assessment) in its surveys.

## B. Differing Approaches

Though consumer-led evaluation teams were found to have many core features in common, there were ways that their work differed. New consumer-led evaluation teams have the opportunity to make choices about how to approach the collection of consumer satisfaction data, based on stakeholder priorities, how the state or local mental health system is set up, organizational capacity, and the expected accountability process.

Approaches of the consumer-led evaluation teams highlighted in this report differed in four significant ways. First, the procedures for conducting interviews and collecting data can be approached quite differently – yet, still result in effective feedback for providers and mental health administrators. Secondly, the purpose for collecting the data can vary. Third, there are different approaches to holding providers, funders, and mental health administrators accountable. Fourth, consumer-led evaluation teams seem to take different approaches regarding whether to have full-time staff, part-time staff, or a combination of the two. The size of their respective budgets varies considerably.

---

[3] Developed by Steven J. Onken, Ph.D., Jeanne M. Dumont, Ph.D., Priscilla Ridgway, M.S.W., Douglas H. Dornan, M.S., and Ruth O. Ralph, Ph.D., http://www.nasmhpd.org/publicationsOTA.cfm.

[4] Developed by the Yale's The Program for Recovery and Community Health (PRCH), http://www.yale.edu/PRCH/tools/rec_selfassessment.html.

## Options for Interviewing and Data Collection Methods

First and foremost, consumer-led evaluation teams value whatever consumers' value. Yet, there are different options for how to engage with consumers and how to collect vital information about their satisfaction with various aspects of the care they are receiving.

One approach is to use unstructured very open-ended interviewing techniques, which are more like conversations. Here, interviewers approach program clients to ask them about any needs or issues they want to discuss with limited prompting from the interviewer. An opening question might be here "How are things going for you here?"; "How do you think about the services you have been getting at this provider?"; "Is the programming meeting your needs?"; "Is there anything you'd like to see added to or changed in the program?" The interviewee guides the interview.

Another option is to use a list of predetermined questions, which can range from open-ended to close-ended. These questions can be broad like above, "If you were in charge of this program, what is the first thing you would change?" Questions can be more focused if you're trying to learn about something more specific from everyone. A few examples include: "What do you think about the psychiatrist on this PACT team?" "What was the admission process like?" "How does the program help you get employment opportunities?"

Close-ended (quantitative) questions require that the respondent choose from a limited number of answer options (e.g,, very satisfied, satisfied, somewhat satisfied, not at all satisfied). Whether open-ended or close-ended questions are used, care is taken to ensure that the issues that matter to consumers are uncovered through the interviewing process. Close-ended questions are much easier to analyze with simple software, but open ended questions can reveal information that a team might not have predicted.

Vital Voices uses mostly close-ended, quantitative question surveys for their interviews but include some open-ended questions. Vital Voices' interviewers ask people to fill out a simple written survey, and then the interviewers go over the survey with the respondent to get more information about satisfaction or dissatisfaction with aspects of their care (i.e. "You are dissatisfied with groups, why is that?"). This method allows the collection of both "hard" numbers for program administrators, but also provides details about why respondents feel the way they do about their care.

The Consumer Satisfaction Team in Philadelphia and the Consumer Quality Team in Maryland use qualitative, open-ended interviews more frequently. Interviewers are trained to facilitate conversations with people about their experiences with services. During the course of the interview, consumers indicate if they are satisfied with the services and if they have needs that are not being met. During training, interviewers learn how to ask questions that are open-ended so they are not leading consumers to answer in a particular way. The CQT concludes interviews by asking 5 quantitative questions taken from the Mental Health Statistics Improvement Program (MHSIP).

For CQI it really depends on the project. CQI often uses very open-ended questions when conducting a needs assessment, and more qualitative questions when evaluating a program that is new and there's not much known about it and kinks probably still need to be worked out. CQI uses close-ended questions for more "established" kinds of services (e.g., inpatient), and if someone answers on the dissatisfaction side of the scale interviewers ask that person to explain why s/he is dissatisfied in his/her own words.

## Purpose of Collecting Data and Reporting

The two primary reasons for these teams to interview consumers are to: 1) evaluate the program, and 2) make sure the respondent's needs are being addressed.

All teams except for CQI focus on making sure individual's needs are addressed. Most interviews are confidential, however, when there is a specific need not being met, the respondent is asked if his/her name can be shared with the agency that can address this need. In some cases this is the provider, but when there is a concern about retribution, it is usually the provider's funder. Accountability reports, meetings, and the analysis of collected data provide information about programs as well as the mental health system.

CQI does not typically advocate for respondents unless there is a rights violation or an abuse. In that case, CQI will follow the approach of the other organizations, and also share relevant contact information to report violations. CQI will also provide clients the contact number of the managed care company if they want to make a complaint directly to it. CQI generally chooses to protect the confidentiality of individuals who participate in surveys. CQI's focus is on providing data on themes heard from a sample of consumers surveyed, rather than details about an individual's specific issue, Therefore, individuals are told that taking part in the survey will not impact their individual services, but will provide feedback to providers about how to improve the quality of care for all clients. CQI believes that it can be most effective at advocating for program and systems improvement overall, as it is statewide and is unable to make site visits on a rotating basis as other consumer-led evaluation teams do.

## Reporting and Accountability Process

The consumer-led evaluation teams share findings with providers and policymakers in a variety of ways. One method for ensuring accountability is to meet directly with providers to share findings. Another approach is to meet with state administrators regularly and then hold them accountable to follow up with providers so that consumers' concerns raised during interviews are addressed. All of the teams do both, either simultaneously or one after the other.

Vital Voices and CQI meet directly with providers to share results, discuss concerns, and brainstorm ways to solve problems. The provider agencies or mental health administrators who fund and oversee provider sites are given a chance to comment on findings, and describe steps they have taken or will take to remediate areas of concern.

Vital Voices sends the report with the summary of consumer and provider findings to the county mental health agency, and then meets with the county to address systematic concerns. Vital Voices later conducts follow-up surveys with consumers to see if providers have addressed their concerns, and they continue to advocate for individuals who reported they were not getting what they needed from the providers.

Consumer Quality Initiatives writes a report detailing their findings and recommendations. CQI then meets with the provider to go over findings, with the managed care staff present. The managed care organization is then responsible for holding providers accountable for making changes to services. The providers are given the opportunity to respond to the findings of the report, and their responses are included in the final report that goes to the managed care organization. Ultimately, accountability follow-up is handled by the managed care organization.

## Staffing Mix

Consumer-led evaluation teams seem to take a variety of approaches regarding how to staff their organizations. CST of Philadeelphia has an entirely full-time staff, all of whom engage in interviewing and contribute to report writing. Consumer Quality Initiatives has a combination of full-time staff who are involved in all aspects of the work, part-time staff who may specialize in one or two tasks (interviewing, administrative work, report writing), as well as several on-call interviewers. Consumer Quality Team of Maryland has both full-time and part-time staff, all of whom engage in interviewing and reporting. CQT has found that part-time work often meets the needs of consumers who are on SSI or SSDI and who are working towards recovery. Vital Voices has a small staff, given its smaller geographic area, who works full or part-time depending on their position.

## Budgets

Consumer-led evaluation teams vary in staff size, scope of work, number of funders, and geographic coverage, therefore their budgets likewise vary. As the organization that covers the smallest geographic region, Vital Voices in Milwaukee had an annual budget of $134,697 in the 2008 fiscal year. Consumer Quality Team in Maryland had a budget of $354,200 in 2008. CQI has a budget of about $500,000, to cover the entire state. Consumer Satisfaction Team, the oldest organization highlighted in this report, had an annual budget of $3,315,336 in 2008. Such a large budget is possible not only because of the significant commitment of public funding agencies in the Philadelphia area to their mission, but also because CST works with multiple consumer populations across a variety of public agencies.

Depending on a variety of factors, consumer-led evaluation team budgets can vary widely. Most organizations start with a modest budget and grow over time, as the organizations hone their approach to consumer satisfaction work, gain increasing respect from state administrators, diversify the scope of their work, and secure contracts from multiple state agencies, universities, or health care organizations. CQI has looked to secure NIH funding for its work.

## C. Innovative Features of Consumer-Led Evaluation Teams

While the consumer-led evaluation teams highlighted in this report have many features in common, each organization has unique approaches to its work. These innovations show the diversity of work that can fall under the rubric of consumer satisfaction.

Highlighted below are "best practices" or innovative features of each of the organizations included in this report.

### Vital Voices, Milwaukee County

*A) A Focus on Consumer Advocacy*

Vital Voices makes a significant commitment to direct consumer advocacy. When Vital Voices formed, they decided that it was not enough to simply report their findings to providers and administrators in hopes that changes would be made. They wanted to make sure that consumers' concerns got addressed. So they made advocacy part of the interview process and started a Consumer Helpline so people did not have to wait for interviewers to show up at their program to make a complaint. The following are some Vital Voices' advocacy success stories:

- Convincing a case manager to communicate with her client via email as much as possible since that was the client's preferred method of communication.

- Connecting a homeless consumer to case management services, which in turn helped her obtain government benefits and other services.

- Ensuring that a client was able to get his anti-psychotic medication after going without it for two days because of social worker error.

The Vital Voices Helpline receives several calls a week from consumers about specific concerns that need to be addressed. Vital Voices staff explains to callers how to self-advocate and they can give referrals, but they also will do direct advocacy for people as needed. Follow-up surveys are done with Helpline callers in order to find out if their needs were addressed. When Vital Voices interviewers conduct interviews, they encourage people to call them when they have concerns.

*B) Follow-up Surveys to Make Sure Needs Addressed*

Vital Voices tries to follow up with each client about complaints, to ask the client whether the issue has been addressed, whether he/she is satisfied with the issue, whether the client needs additional advocacy and whether the client is satisfied with Vital Voices' efforts. According to follow-up surveys, 90-95% of consumers' complaints that Vital Voices follows up on with providers are eventually addressed.

*C) Consumer/Family Members Administering the MHSIP Survey*

Vital Voices administers the Mental Health Statistics Improvement Program[5] (MHSIP) survey annually. The MHSIP survey provides a way for providers to evaluate how they are doing in comparison to other organizations and to their own performance over time. Each fall, Vital Voices administers MHSIP surveys to approximately 500 people at community-based agencies. The county then analyzes the data and gives results to agencies so that concerns can be addressed. Vital Voices receives copies of the reports and then uses the results to figure out what areas of consumer satisfaction and quality of care need addressing. One of the areas that frequently comes up through the MHSIP is insufficient consumer input into treatment planning. With that information at the program level, Vital Voices can then find out why that is. The MHSIP, on its own, does not explain why. Yet, the MHSIP is administered every year to the same programs, which allows Vital Voices and the county to see trends over time.

## Consumer Quality Initiatives, Massachusetts

*A) Data-Driven Reports*

CQI's satisfaction surveys focus on generating quantitative data-driven reports for policy makers and providers that have the potential to have a broader impact on system change, beyond addressing individuals' and even individual program needs. The results of these surveys generate statistics at the provider-level, and also analyze across providers how people in a number of hospitals have answered the questions, or, at times, how a question has been answered across provider levels of care. Thus, CQI is able to present to trade groups (hospital, outpatient) the aggregate responses of, for example, 300 consumers at 20 provider sites, demonstrating industry trends.

*B) Expanding from Consumer Satisfaction into Community-based Participatory Action Research*

Early on CQI's executive director recognized that it would have little impact on system reform by only helping currently funded programs to improve. It thus aimed to go right to the consumers outside of the program context to ask what research they wanted to see, An early example was the community's recommendation to look at the needs of youth aging out of mental health services. CQI hired and trained several young adults to participate with them in developing an open-ended questionnaire, interviewing young adults across the state, and then writing a report, which recommended a complete overhaul of the little that was being done. The key was that the young adults made the presentation to DMH staff and stakeholders, and thus many of those recommendations have taken hold over the years.

---

[5] http://www.mhsip.org/whatis.html

### *C) Emphasis on building community capacity for research and evaluation*

Realizing that it could not conduct all the important consumer-driven research/evaluation that needed to be achieved, CQI developed a researcher training curriculum for consumers, and has several contracts to conduct that training, while engaging in participatory research with the trainees. In addition, CQI found that universities, which obtain a large amount of research funding, were not warm to consumer participation in research. CQI began to advocate for greater involvement, and now consults with several universities on the meaningful involvement of consumers in the production of research. In addition, CQI's Executive Director is currently one of two principal investigators on a National Institutes of Mental Health funded community-based participatory action research project. Following up on a previous grant, a team of consumers will hired to be community research associates, go through an extensive training on how to conduct research, and then participate in any part of the research process that they choose.

## The Consumer Satisfaction Team, Philadelphia

### *A) Encompassing Reach Across Provider Type and Governmental Agencies*

As the first consumer evaluation team in the United States, CST has continued to expand its reach over time across most levels of care and several government agencies related to behavioral health care. CST staff makes unannounced visits every day of the week to a large variety of program sites including:

- day treatment programs,
- drop-in centers, clubhouses,
- detoxification facilities,
- rehabilitation and recovery programs,
- outpatient clinics,
- inpatient facilities,
- crisis centers, as well as
- children's residential treatment facilities and schools.

With community-based services, CST interviews consumers in their places of residence. In addition to regularly scheduled visits to these provider types, CST also takes on special projects to look at specific service types, such as crisis response centers, partial hospitalization programs, therapeutic summer camps and case management services for consumers. The special projects allow for the collection of consumer perspectives around an identified problem area.

Over the years, CST has also expanded to include provider sites under the purview of multiple governmental agencies. Its annual contract began with the Philadelphia Office of Mental Health under the Department of Behavioral Health, but has expanded to include the Coordinating Office for Drug and Alcohol Abuse Programs, the Delaware County Office of Mental Health, and the Philadelphia Department of Human Services.

### *B) Interview Team Staffing Approach*

Because CST interviews across a wide variety of provider types, CST's large staff is organized into teams based on governmental agency/service or population focus: Drug and Alcohol Team, Department of Human Services Team, Family Team, and Mental Health Team. Everyone in the organization, including the Executive Director, is an interviewer. Each team has a team leader who coordinates interview schedules and oversees the work of that team. The team based approach allows CST to staff each team with consumers or family members who have experience with that service type. In addition to hiring staff with experience with the mental health system, CST staff now includes people with experience in substance abuse, child and adolescent services, and former and current clients of Department of Human Services.

### *C) On-going Family Member Survey*

Consumer Satisfaction Team regularly solicits the input of family members during visits to provider sites, if family members are present, and receives and documents phone calls from family members. In addition to talking directly with family members in-person or by phone, CST conducts an ongoing survey to ascertain family members' perspectives on their loved one's services.

## Consumer Quality Team of Maryland

### A) *Balancing Staff's Individual Needs with Team Building*

Because CQT is a start-up team, particular emphasis has been placed on providing regular training and support to staff. CQT's personnel policies reflect a commitment to employee wellness and provide options and supports for employees dealing with a current mental health issue. In addition, CQT has on-call interviewers who can be available if a staff member needs to take a leave of absence.

CQT also has a training and supervision process that is supportive of individual needs and is focused on team building. They use a training manual, but the personal training period varies widely depending on the employee's background. For some people, interviewing is a totally new activity or they may have limited knowledge about some areas of the public mental health system, so their training may last 2-3 months. Others are more experienced and can be trained in as little as 5 weeks. Training is tailored to support and encourage each employee's professional growth.

To support team building and ongoing training, every Friday is an in-office day so staff can finish up reports, plan the next week and do training activities. Role-playing and roundtable discussions about what came up during the past week are utilized. There is a focus on collaboration and team-teaching; individual staff members and teams share challenges that they faced during the week and all team members participate in brainstorming how to improve.

*B) Outside Assessment of Consumer Evaluation Work*

As the mental health system continues to move towards evidence-based practices, the decision was made to have the work of Consumer Quality Team of Maryland evaluated by an outside agency from the beginning. The first year evaluation focused on staff member perspectives on the organizational start-up process. The second year evaluation was broader and included the perspectives of consumers, providers, advisory board members, and Core Service Agency staff.

The outside evaluators used a variety of methods to collect feedback, including phone interviews, surveys and focus groups. The information collected is used to change and improve the evaluation process. For example, interviews with CQT staff revealed that they needed something to clearly identify them when they went to a site to interview consumers. The organization purchased badges and polo shirts that clearly identified them as CQT staff members. In response to consumer feedback, a toll-free number was obtained and added to the promotional materials. In response to provider feedback, reports were changed to provide an executive summary of the report. In addition to critical information used for improvements, the evaluation allows Consumer Quality Team staff to hear positive feedback from stakeholders and to better document the ways their work is making a difference.

*C) One Consumer Evaluation Program for the State*

Maryland is one of the smaller states, in terms of its geographic area. The CQT began as a pilot program in three jurisdictions. After just six months of operation, the CQT oversight committee, with the support of the providers in the pilot jurisdiction, asked for and received from the Mental Hygiene Administration funding to take the program statewide. The decision was made to locate the CQT office in central Maryland (Baltimore), and to expand the program out from the pilot area. When the more distant areas of the state are reached, the teams will stay overnight in the area, visiting multiple locations. This will not only minimize the program's administrative costs, it will also ensure fidelity to the program and to the collection of comparable data.

## D. Future Directions and Expanding Impact of Consumer-Led Evaluation Teams

Consumer-led evaluation teams are looking to the future with an eye towards how to expand their impact. Two themes arose about expanding the influence of the organizations. Three out of the four organizations were working to expand their work to other populations and provider types.

- Consumer Quality Team of Maryland has expanded its adult program visits from 3 jurisdictions to 14 jurisdictions, and plans to cover all 23 jurisdictions. In addition, plans exist to then begin visiting the child and adolescent mental health programs.

- Vital Voices is planning to branch out into consumer evaluation of services for alcohol and other drug abuse (AODA) services, and hopes to evaluate newly developed housing programs for people with mental illness and/or a history of homelessness.
- Consumer Quality Initiatives of Massachusetts continues to expand its work to both addiction programs and services for people with physical disabilities. Within mental health, CQI continues to evaluate and research peer run and peer specialist[6] services. In addition, it has been working to establish a participatory action research center.

Expansion often requires an increase in the size of the staff and greater administrative capacity (e.g., accounting, contract management). It also means that other kinds of work expertise will need to be brought in, including consumer staff of non-mental health services. Business or organizational consultants can often be useful here.

Increased funding will either be secured by contracting with other health agencies or through subcontracts with other research and evaluation organizations. For instance, the Consumer Satisfaction Team of Philadelphia has expanded by contracting for consumer satisfaction work with the Department of Human Services. Other consumer-led evaluation teams are contracting with large providers for more focused projects, or with other research or evaluation organizations that might bring different skills and capacities to consumer satisfaction work.

## Recommendations for Creating and Sustaining a Consumer-Led Evaluation Team

Based on the findings presented above about the core features of consumer-led evaluation teams, differing approaches, and innovative features of the four organizations highlighted in this report, we have developed a set of recommendations for those planning to start a consumer-led evaluation team in their own community.

1. When starting a consumer-led evaluation team, establish its mission and goals. One issue is whether a team plans to focus on evaluating systems and programs, and/or responding to clients' personal issues. The idea is to fill an unmet need(s)

   Regardless, the ultimate goals for these teams is to see that providers and (quasi-)governmental funders are held accountable for addressing the findings and reports of the team. An effective accountability process requires mutual respect and regular communication.

2. Start small and get it right. In the beginning, focus on a particular provider type or limited geographic area. Over time expand out to a larger geographical area or to other service delivery areas, such as child and adolescent services.

---

[6] Peer specialists are people in recovery who are trained, and increasingly certified, to provide paid peer-to-peer support at a variety of mental health programs.

3. Develop a strong, collaborative relationship with both the agencies who are in a position to hold providers accountable (e.g., government, managed care) and the providers themselves. This buy-in is especially important since the providers often facilitate the recruitment of consumers to be interviewed.

4. Determine how you want to collect consumer perspectives and spend time honing your methods for collecting consumer satisfaction information. Discuss the pros and cons for different types of data collection approaches (i.e. open-ended versus close-ended questions; consumer comments verbatim versus statistics). It may be that the method you use depends on the circumstances.

5. Think through data management and reporting procedures. Consider what types of information providers and other stakeholders will value and what kind of data will move them to action. Then, create an organized and efficient method for moving from data collection to data storage to analysis and reporting.

6. Carefully think through how you can protect the confidentiality of the responses of the people you interview.

7. Provide people hired with the appropriate level of supervision, training and support. Meet the specific vocational needs of individuals. Some staff may need "reasonable accommodations"[7] so that they can perform their "essential job functions." Consumers working for consumer-led evaluation teams are a powerful reminder to providers and other consumers that recovery is real.

8. Have an effective training program for new staff. Determine whether you want staff to specialize in particular tasks (interviewing, writing reports, meeting with providers) or to be "generalists", who participate in all or most phases of the data collection and reporting process. Consumer satisfaction work can require many skills – interpersonal, time management, analytic thinking, writing, and presenting information.

9. Once the organization is established and policies and procedures become solidified, consider how to diversify your funding portfolio. For example, look for opportunities to do consumer satisfaction work for other governmental agencies, providers, consumer groups and universities.

---

[7] Per federal (Americans with Disabilities Act) and sometimes state law

# MFP QUALITY OF LIFE SURVEY
# DRAFT

## RESPONDENT INFORMATION

Respondent Name: ________________________________________

Respondent Street Address: ________________________________________

Respondent City: ________________________________________

Respondent State: ________________________________________

Respondent ZIP Code: ________________________________________

Social Security Number: ________________________________________

Medicaid ID number: ________________________________________

☐ *Check here if the Sample Member is deceased and record date of death:*

[__________] [__________] [___________] ➨ **GO TO END**
Month Day Year

Hello, my name is _______ and I am from ________. I'm here to ask for your help with an important study of Medicaid beneficiaries in the state of __________. The Quality of Life Survey, sponsored by the Centers for Medicare & Medicaid Services (CMS) and the state of __________, is an essential part of an evaluation of the Money Follows the Person Program, a program designed to help Medicaid beneficiaries transition out of institutional care into the community. I'd like to ask you some questions about your housing, access to care, community involvement, and your health and well-being. Results from the study will help CMS and the state of __________ evaluate how well its programs are meeting the needs of Medicaid beneficiaries like you.

Before we begin, let me assure you that all information collected will be kept strictly confidential and will not be reported in any way that identifies you personally. Your answers will be combined with the answers of others and reported in such a way that no single individual could ever be identified. Further, the information collected will not be used by anyone to determine your continuing eligibility for Medicaid benefits. We are collecting this information for research purposes only. However, I may be required to report any instances of abuse or neglect that you tell me about to authorities. Your participation is completely voluntary and if we come to any question you prefer not to answer, just tell me and we'll move on to the next one.

If you have any questions, please stop me and ask me. Also, please let me know if you do not understand a question or if you would like me to repeat it.

## MODULE 1: LIVING SITUATION

1. I'm going to ask you a few questions about the place you live. About how long have you lived (here/in your home)?

Probe: Your best estimate is fine.

*Interviewer: If respondent indicates less than 1 month, enter 1 month.*

[_________] Years [__________] Months ➡ **GO TO QUESTION 2**

DON'T KNOW ...................................... DK
REFUSED ............................................ R

1a. Would you say you have lived here more than five years?

Yes ............................................................. 01
No ............................................................... 02
Don't Know .................................................. DK
Refused ....................................................... R

2. *Interviewer: Does sample member live in a group home or nursing facility?*

Yes ......................................................................... 01
No ........................................................................... 02
Don't Know .............................................................. DK
Refused ................................................................. R

3. Do you like where you live?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| DON’T KNOW | DK |
| REFUSED | R |

4. Did you help pick (this/that) place to live?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON’T KNOW | DK |
| REFUSED | R |

5. Do you feel safe living (here/there)?

| | | |
|---|---|---|
| Yes | 01 | ➨ **GO TO QUESTION 6** |
| No | 02 | |
| DON’T KNOW | DK | ➨ **GO TO QUESTION 6** |
| REFUSED | R | ➨ **GO TO QUESTION 6** |

5a. How often do you feel unsafe living (here/there)?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON’T KNOW | DK |
| REFUSED | R |

6. Can you get the sleep you need without noises or other disturbances where you live?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| DON’T KNOW | DK |
| REFUSED | R |

## MODULE 2: CHOICE AND CONTROL

7. Can you go to bed when you want?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| DON’T KNOW | DK |
| REFUSED | R |

8. Can you be by yourself when you want to?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| DON'T KNOW | DK |
| REFUSED | R |

9. When you are at home, can you eat when you want to?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| DON'T KNOW | DK |
| REFUSED | R |

10. Can you choose the foods that you eat?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| DON'T KNOW | DK |
| REFUSED | R |

11. Can you talk on the telephone without someone listening in?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| No access to telephone | 04 |
| DON'T KNOW | DK |
| REFUSED | R |

12. Can you watch TV when you want to?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| No access to TV | 04 |
| DON'T KNOW | DK |
| REFUSED | R |

13. **[AFTER TRANSITION ONLY]** Some people get an allowance from the state to pay for the help or equipment they need. Do you get an allowance like this?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 14** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 14** |
| REFUSED | R | ➡ **GO TO QUESTION 14** |

13a. **[AFTER TRANSITION ONLY]** In the last 12 months, what help or equipment did you buy with this allowance?

**[Code all that apply]**

| | |
|---|---|
| Modified Home | 01 |
| Modified Car | 02 |
| Special Equipment | 03 |
| Paid Help | 04 |
| Transportation | 05 |
| Household Goods | 06 |
| Security Deposit | 07 |
| Other | 08 |
| DON'T KNOW | DK |
| REFUSED | R |

## MODULE 3: ACCESS TO PERSONAL CARE

14. Now I'd like to ask you about some everyday activities, like getting dressed or taking a bath. Some people have no problem doing these things by themselves. Other people need somebody to help them. First, does anyone help you with things like bathing, dressing, or preparing meals?

Probe: Please include any help received by another person, including cueing or standby assistance.

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 15** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 15** |
| REFUSED | R | ➡ **GO TO QUESTION 15** |

14a. Do any of these people get paid to help you?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 15** |
| Don't Know | DK | ➡ **GO TO QUESTION 15** |
| Refused | R | ➡ **GO TO QUESTION 15** |

14b. Do you pick the people who are paid to help you?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Don't Know | DK |
| Refused | R |

15. Do you ever go without a bath or shower when you need one?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 16** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 16** |
| REFUSED | R | ➡ **GO TO QUESTION 16** |

15a. How often do you go without a bath or shower when you need one? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

15b. Is this because there is no one there to help you?

Probe: Please include any help received by another person, including cueing or standby assistance.

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

16. Do you ever go without a meal when you need one?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 17** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 17** |
| REFUSED | R | ➡ **GO TO QUESTION 17** |

16a. How often do you go without a meal when you need one? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

16b. Is this because there is no one there to help you?

Probe: Please include any help received by another person, including cueing or standby assistance.

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

17. Do you ever go without taking your medicine when you need it?

Probes: Medicines are pills or liquids that are given to you by a doctor to help you feel better.

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 18** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 18** |
| REFUSED | R | ➡ **GO TO QUESTION 18** |

17a. How often do you go without taking your medicine when you need it? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

17b. Is this because there is no one there to help you?

Probe: Please include any help received by another person, including cueing or standby assistance.

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

18. Are you ever unable to use the bathroom when you need to?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 19** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 19** |
| REFUSED | R | ➡ **GO TO QUESTION 19** |

18a. How often are you unable to use the bathroom when you need to? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

18b. Is this because there is no one there to help you?

Probe: Please include any help received by another person, including cueing or standby assistance.

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

19. **[AFTER TRANSITION ONLY]** Have you ever talked with a case manager or support coordinator about any special equipment or changes to your home that might make your life easier?

Probe: Equipment means things like wheelchairs, canes, vans with lifts, and automatic door opener.

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 20** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 20** |
| Not Applicable | N/A | ➡ **GO TO QUESTION 20** |
| REFUSED | R | ➡ **GO TO QUESTION 20** |

19a. **[AFTER TRANSITION ONLY]** What equipment or changes did you talk about?

________________________________________

________________________________________

________________________________________

________________________________________

| | |
|---|---|
| DON'T KNOW | DK |
| REFUSED | R |

19b. **[AFTER TRANSITION ONLY]** Did you get the equipment or make the changes you needed?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| In Process | 03 |
| DON'T KNOW | DK |
| REFUSED | R |

20. **[AFTER TRANSITION ONLY]** Please think about all the help you received during the last week *around the house* like cooking or cleaning. Do you need *more* help with things around the house than you are now receiving?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

21. **[AFTER TRANSITION ONLY]** During the last week, did any family member or friends help you with things around the house?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 22** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 22** |
| REFUSED | R | ➡ **GO TO QUESTION 22** |

21a. **[AFTER TRANSITION ONLY]** Please think about *all* the family members and friends who help you. About how many hours did they spend helping you yesterday?

Probe: Your best estimate is fine.

*Interviewer: if less than one hour, enter 1 hour.*

[_________]
Hours

| | |
|---|---|
| DON'T KNOW | DK |
| REFUSED | R |

## MODULE 4: RESPECT AND DIGNITY

Note: If Q14 = No, DK or R ➡ **GO TO QUESTION 27**

*Interviewer: For questions in this module, refer to your state's policy on reporting any suspected incidents of abuse and neglect. For this survey, record only reports of current abuse.*

22. You said that you have people who help you. Do the people who help you treat you the way you want them to?

| | | |
|---|---|---|
| Yes | 01 | ➡ **GO TO QUESTION 23** |
| No | 02 | |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 23** |
| REFUSED | R | ➡ **GO TO QUESTION 23** |

22a. How often do they not treat you the way you want them to? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

23. Do the people who help you listen carefully to what you ask them to do?

| | | |
|---|---|---|
| Yes | 01 | ➡ **GO TO QUESTION 24** |
| No | 02 | |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 24** |
| REFUSED | R | ➡ **GO TO QUESTION 24** |

23a. How often do they not listen to you? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

24. **[Optional]** Have you ever been physically hurt by any of the people who help you now?

Probe: Physically hurt means someone could have pushed, kicked, or slapped you.

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 25** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 25** |
| REFUSED | R | ➡ **GO TO QUESTION 25** |

24a. **[Optional]** What happened when the people who help you now physically hurt you?

________________________________________

________________________________________

________________________________________

________________________________________

| | |
|---|---|
| DON'T KNOW | DK |
| REFUSED | R |

24b. **[Optional]** How many times have you been physically hurt by the people who help you now?

Probe: Your best guess is fine.

[__________]
Times

| | |
|---|---|
| DON'T KNOW | DK |
| REFUSED | R |

25. **[Optional]** Are any of the people who help you now mean to you or do they yell at you?

Probe: Do they treat you in a way that makes you feel bad or do they hurt your feelings?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 26** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 26** |
| REFUSED | R | ➡ **GO TO QUESTION 26** |

25a. **[Optional**] How often are they mean to you? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

26. **[Optional]** Have any of the people who help you now ever taken your money or things without asking first?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➨ **GO TO QUESTION 27** |
| DON'T KNOW | DK | ➨ **GO TO QUESTION 27** |
| REFUSED | R | ➨ **GO TO QUESTION 27** |

26a. **[Optional]** How many times have they taken your money or things without asking first?

Probe: Your best guess is fine.

[__________]
Times

| | |
|---|---|
| DON'T KNOW | DK |
| REFUSED | R |

## MODULE 5: COMMUNITY INTEGRATION AND INCLUSION

27. I'd like to ask you a few questions about things you do. Can you see your friends and family when you want to see them?

*Interviewer: Code "yes" if respondent indicates that they have either gone to see friends or family or that friends and family have come to visit them.*

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➨ **GO TO QUESTION 28** |
| DON'T KNOW | DK | ➨ **GO TO QUESTION 28** |
| REFUSED | R | ➨ **GO TO QUESTION 28** |

27a. How often do you see your friends and family when you want to see them? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

28. Can you get to the places you need to go, like work, shopping, or the doctor's office?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➨ **GO TO QUESTION 29** |
| DON'T KNOW | DK | ➨ **GO TO QUESTION 29** |
| REFUSED | R | ➨ **GO TO QUESTION 29** |

28a. How often do you get to the places you need to go, like work, shopping, or the doctor's office? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

29. Is there anything you *want* to do outside [the facility/your home] that you can't do now?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➨ **GO TO QUESTION 30** |
| DON'T KNOW | DK | ➨ **GO TO QUESTION 30** |
| REFUSED | R | ➨ **GO TO QUESTION 30** |

29a. What would you like to do that you don't do now?

______________________________________

______________________________________

______________________________________

______________________________________

| | |
|---|---|
| DON'T KNOW | DK |
| REFUSED | R |

29b. What do you need to do these things?

______________________________________

______________________________________

______________________________________

______________________________________

| | |
|---|---|
| DON'T KNOW | DK |
| REFUSED | R |

30. When you go out, can you go by yourself or do you need help?

| | | |
|---|---|---|
| Go out Independently | 01 | ➡ **GO TO QUESTION 31** |
| Need Help | 02 | |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 31** |
| REFUSED | R | ➡ **GO TO QUESTION 31** |

30a. Please think about *all* the help you received during the last week with *getting around the community,* such as shopping and going to a doctor's appointment, do you need *more* help getting around than you are receiving?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

31. **[AFTER TRANSITION ONLY]** Are you working for pay right now?

Probe: Do you get any money for doing work?

| | | |
|---|---|---|
| Yes | 01 | ➡ **GO TO QUESTION 32** |
| No | 02 | |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 32** |
| REFUSED | R | ➡ **GO TO QUESTION 32** |

31a. **[AFTER TRANSITION ONLY]** Do you want to work for pay?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

32. **[AFTER TRANSITION ONLY]** Are you doing volunteer work or working without getting paid?

Probe: Are you doing work but not getting any money for it?

| | | |
|---|---|---|
| Yes | 01 | ➡ **GO TO QUESTION 33** |
| No | 02 | |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 33** |
| REFUSED | R | ➡ **GO TO QUESTION 33** |

32a. **[AFTER TRANSITION ONLY]** Would you like to do volunteer work or work without getting paid?

Probe: would you like to do work without getting paid for it?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

33. I'd like to ask you a few questions about how you get around. Do you go out to do fun things in your community?

Probe: These are things that you enjoy such as going to church, the movies or shopping?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

34. When you want to go somewhere, can you just go, do you have to make some arrangements, or do you have to plan many days ahead and ask people for help?

| | |
|---|---|
| Decide and Go | 01 |
| Plan Some | 02 |
| Plan Many Days Ahead | 03 |
| DON'T KNOW | DK |
| REFUSED | R |
| N/A | NA |

35. Do you miss things or have to change plans because you don't have a way to get around easily?

Probe: Do you have to miss things because it is hard for you to get there?

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| Sometimes | 03 |
| DON'T KNOW | DK |
| REFUSED | R |

36. Is their any medical care, such as a medical treatment or doctor's visits, which you have not received or could not get to within the past month?

Probe: The medical care includes doctor visits or medical treatments that you may need.

| | |
|---|---|
| Yes | 01 |
| No | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

## MODULE 6: SATISFACTION

37. Taking everything into consideration, during the past week have you been happy or unhappy with the help you get with things around the house or getting around your community?

| | | |
|---|---|---|
| Happy | 01 | **➡ GO TO QUESTION 37a** |
| Unhappy | 02 | **➡ GO TO QUESTION 37b** |
| DON'T KNOW | DK | **➡ GO TO QUESTION 38** |
| REFUSED | R | **➡ GO TO QUESTION 38** |

37a Would you say you are a little happy or very happy?

| | | |
|---|---|---|
| A little happy | 01 | **➡ GO TO QUESTION 38** |
| Very happy | 02 | **➡ GO TO QUESTION 38** |
| Don't Know | DK | **➡ GO TO QUESTION 38** |
| Refused | R | **➡ GO TO QUESTION 38** |

37b Would you say you are a little unhappy or very unhappy?

| | |
|---|---|
| A little unhappy | 01 |
| Very unhappy | 02 |
| Don't Know | DK |
| Refused | R |

38. Taking everything into consideration, during the past week have you been happy or unhappy with the way you live your life?

| | | |
|---|---|---|
| Happy | 01 | **➡ GO TO QUESTION 38a** |
| Unhappy | 02 | **➡ GO TO QUESTION 38b** |
| DON'T KNOW | DK | **➡ GO TO QUESTION 39** |
| REFUSED | R | **➡ GO TO QUESTION 39** |

38a. Would you say you are a little happy or very happy?

| | | |
|---|---|---|
| A little happy | 01 | **➡ GO TO QUESTION 39** |
| Very happy | 02 | **➡ GO TO QUESTION 39** |
| Don't Know | DK | **➡ GO TO QUESTION 39** |
| Refused | R | **➡ GO TO QUESTION 39** |

38b. Would you say you are a little unhappy or very unhappy?

| | |
|---|---|
| A little unhappy | 01 |
| Very unhappy | 02 |
| Don't Know | DK |
| Refused | R |

## MODULE 7:  HEALTH STATUS

39. During the past week have you felt sad or blue?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 40** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 40** |
| REFUSED | R | ➡ **GO TO QUESTION 40** |

39a. How often have you felt sad and blue? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

40. During the past week have you felt irritable?

Probe: Irritable means grumpy or easily upset about things in your life.

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 41** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 41** |
| REFUSED | R | ➡ **GO TO QUESTION 41** |

40a. How often have you felt irritable? Would you say only sometimes or most of the time?

Probe: Irritable means grumpy or easily upset about things in your life.

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

41. During the past week have you had aches and pains?

| | | |
|---|---|---|
| Yes | 01 | |
| No | 02 | ➡ **GO TO QUESTION 42** |
| DON'T KNOW | DK | ➡ **GO TO QUESTION 42** |
| REFUSED | R | ➡ **GO TO QUESTION 42** |

41a. How often do you have aches and pain? Would you say only sometimes or most of the time?

| | |
|---|---|
| Sometimes | 01 |
| Most of the Time | 02 |
| DON'T KNOW | DK |
| REFUSED | R |

## CLOSEOUT

42. Those are all the questions I have you now. We would like to talk with you in about a year or so to find out how you are doing. In case we have trouble reaching you, what is the name, address, and phone number of a close relative or friend who is not living with you and is likely to know your location in the future? For example, a mother, father, brother, sister, aunt, uncle, or close friend.

| | | |
|---|---|---|
| No Contact Available............................. | 01 | ➡ **GO TO QUESTION 43** |
| Contact Available .................................. | 02 | |

42a. Contact Name: ________________________________

42b. Contact Street Address: ________________________________

42c. Contact City: ________________________________

42d. Contact State: ________________________________

42e. Contact ZIP ________________________________

42f. Contact Phone: ________________________________

43. *Interviewer: Did you complete the interview with the sample member alone, the sample member who was assisted by another, or with a proxy?*

| | |
|---|---|
| Sample Member Alone........................................... | 01 |
| Sample Member with Assistance............................ | 02 |
| Proxy ...................................................................... | 03 |

44. *Interviewer: Record date the interview was completed:*

[__________] [__________] [___________]
Month Day Year

➡ **END INTERVIEW**