MAYO CLINIC

Harmonizing and Consolidating the Measurement of Patient-Reported Outcomes at Mayo Clinic

A Position Statement for the Center for the Science of Health Care Delivery

Prepared by:

David T. Eton, Ph.D.
Timothy J. Beebe, Ph.D.
Philip T. Hagen, M.D.
Michele Y. Halyard, M.D.
Victor M. Montori, M.D.
James M. Naessens, Sc.D.
Jeff A. Sloan, Ph.D.
Carrie A. Thompson, M.D.
Douglas L. Wood, M.D.

December 6, 2012

Table of Contents

Executive Summary

A patient-centered institution must be aware of the experience patients have of health and health care. Patient-Reported Outcomes (PROs) capture that experience and complement disease-centered clinical measures to inform practice, assess the value of health care, and facilitate patient-centered research. This *Position Statement* is written primarily for Mayo Clinic clinicians, clinical teams, and administrators. It is intended to be used as a tool to educate and inform these stakeholders about the merits of collecting and reporting PROs as well as the importance of consolidating their measurement at Mayo Clinic. The statement is divided into six chapters each written by a content expert or team of experts in a particular area. A glossary of PRO-related terms is also provided in an appendix. The writing team consisted of individuals with expertise in the following areas: patient-reported outcome / quality of life (PRO-QOL) science, survey methods, health-care policy research, health-care delivery research, value/quality measurement, preventive medicine, and clinical practice (including practice re-design).

To encapsulate our principal findings and lessons learned, we provide 2 to 3 "take away" messages from each chapter.

<u>Chapter 1: Introduction to PROs</u>

1. A PRO is any report of the status of a patient's health condition (e.g., symptoms, functional status), health behavior, or experience with health care or the health-care system that comes *directly from the patient* without interpretation by a clinician or anyone else.

2. The utility of PROs in clinical research (e.g., comparative effectiveness research, outcomes research) is well documented. However, today there is increasing interest and great potential in using *individual-level* PRO data to improve patient care and *aggregate-level* PRO data for the purposes of accountability and performance improvement of health-care providers.

<u>Chapter 2: Implementing PROs into practice</u>

1. The use of PRO-QOL data in clinical practice has the potential to enhance patient care by highlighting issues of concern, increasing shared decision making, and increasing satisfaction with care.

2. PRO-QOL data collection can be applied to not only care that happens in the traditional office setting, but also health care from afar such as telehealth via use of electronic means (ePRO).

3. Although not always easy to implement, examples of effective use of PRO-QOL data are available to guide clinicians in this endeavor.

<u>Chapter 3: What is Mayo currently doing / not doing?</u>

1. Initiatives that exist are all fairly *embryonic*, and mostly *motivated by compliance* with accreditation, public reporting, or payer organizations. Either they are in the pilot states of

data collection and analysis, or they have robust data collection, but the data are used in a limited way in guiding clinical care, or in analyzing the practice. There is no Mayo Clinic strategy for PROs.

2. There is no one group that is coordinating the PRO design, selection, collection, data management, interpretation, and integration into clinical pathways, and there is extreme paucity or limited access to performance or outcome data.

Chapter 4: Capturing overall (subjective) patient functioning

1. In the context of PROs, overall (general) patient functioning can be defined in terms of four inter-related concepts – symptoms, functional status, general health perception, and overall quality of life.

2. The PROMIS Global-10 measure would appear to be an excellent candidate for capturing overall general health status in Mayo Clinic patients.

Chapter 5: Next steps in PRO roll-out at Mayo Clinic

1. While there are some success stories of clinically-integrated and efficacious PRO collection at Mayo, much can be learned from the experiences of our contemporaries in this realm, particularly Dartmouth and Partners Healthcare (Harvard).

2. Robust and continued engagement with key internal stakeholders such as patients and the Practice Convergence Council will be imperative for the success of any PRO data collection system at Mayo.

3. Mayo needs to consider a range of data collection vehicles and venues in the design of its PRO infrastructure, indexing the identification of these off the preferences of the patient.

Chapter 6: How do we influence others?

1. Being the “ultimate” patient-centered measures, PROs will provide key measures in determining the value of health care.

2. Engaging multiple stakeholders through multiple means will encourage the adoption of PROs and demonstrate their value.

3. Mayo Clinic needs to develop a robust PRO system before it has the recognition and credibility to influence others.

Mayo Clinic and other institutions have experience with PROs that should be harnessed to launch a comprehensive initiative, one that unites the input of clinicians and patients. The Center for the Science of Health Care Delivery is poised to support the institution in its journey toward the universal and ubiquitous collection of PROs and their use for practice, to assess the value of health care and for research.

Chapter 1
Patient-Reported Outcomes – Introduction and background

By David T. Eton, Ph.D.

What is a Patient-Reported Outcome (PRO)?

A patient-reported outcome or PRO is any report of the status of a patient's health condition, health behavior, or experience with health care that comes *directly from the patient* without interpretation by a clinician or anyone else.[1] The term "PRO" is an umbrella that classifies a range of different patient-related concepts. This includes personal reports of health status such as assessments of functional status, symptoms, and "health-related quality of life" (HRQL). It also includes assessments of health-related behaviors, both those that are detrimental to health (e.g., smoking) and those that promote good health (e.g., exercise). Although health behavior assessments are often considered predictors of health outcome, they can also be construed as health outcomes inasmuch as health-care interventions can have an impact on them.[1] Finally, PROs can also be used to capture the patient's experience with health care in the form of assessments of patient satisfaction (e.g., treatment satisfaction), patient activation (e.g., shared decision making), and consumer experience with health-care services (i.e., quality of care).[1] Not all medical or health information collected from patients constitutes a PRO. For example, demographic characteristics, current medication lists, and personal and family medical history are all important pieces of health information that a patient may provide. However, this information does not represent a *health outcome* per se and is therefore not a PRO.

PRO information can be obtained from patients in a variety of ways including interviews, questionnaires (paper, dedicated electronic device, web, or phone - interactive voice response), or diaries. PROs are distinct from measures of clinical efficacy such as laboratory and biometric measurements in that they are informed entirely by the subjective impressions of patients, without interpretation by any other person. While physiological measures are useful markers of patient health status, they fail to consider the patient's perspective. Laboratory and biometric measures cannot incorporate aspects of health that are important to the patient such as overall functional capacity, well-being, and health behavior. Most importantly, perceptions of patient well-being by individuals other than the patient can be markedly different and even uncorrelated to what the patient perceives and holds important. Hence, the inclusion and routine collection of PROs *in addition to* traditional clinical endpoints can provide a more complete picture of how a patient is functioning and ultimately facilitate the path to a more patient-centric care that respects the needs, values, and preferences of each individual.

What properties does a good PRO have?

A PRO is assessed using a "measure" or "tool." To alleviate confusion, throughout this manuscript the term PRO will be used interchangeably to refer to both the health concept being assessed and the tool used to measure it. The following represent properties of a "good" PRO:

- **It is simple.** As a general rule of thumb, a PRO measure should not require reading skills beyond that of a 12-year-old (6th grade level).[2] Questions and items should contain no jargon or other terms that a patient might not be familiar. Rating formats (e.g., scales) should be intuitive and instructions should be easy to understand.

- **It is brief.** Question sets that are too long will not be completed. The more questions that a respondent is asked to answer, the less likely they are to answer them.[3] As a general rule of thumb, any question set that requires more than about 12-15 minutes to complete is probably too long. Brevity is essential when PROs are to be used to inform clinical care or improve performance of health-care services.

- **It is informed by patients.** PROs developed using direct input from patients tend to be more relevant and meaningful to patients than those that rely on clinician and researcher input alone.

- **It is reliable and valid.** A reliable PRO yields responses that are reproducible and consistent (e.g., stable within a brief period). A valid PRO is one that measures what it is intended to measure and not something else. There are many different types of validity. The importance of each type will depend on the use of the PRO. For instance, when a PRO is intended for use in clinical practice, face, content, and predictive validity are the most important to demonstrate. Face validity indicates whether, on the face of it, the questions on a PRO look like they are measuring what they are supposed to be measuring. Content validity, a related concept, is a judgment of whether the PRO samples the most relevant and important content that it was designed to measure. Predictive validity is the utility of a PRO measure to predict some future outcome.

- **It is responsive to change.** PRO measures for prospective use should be able to detect underlying changes in the experience of patients over time. For example, changes in patient-reported health status should coincide with objective changes in patient health.

- **It is easily scored and interpreted, and predicates clinical action.** Responses to PRO measures (be they single-item assessments or multi-item scales) should be easy to score, interpret and use. Measures that can be immediately scored and do not require complicated scoring algorithms or technical assistance from third parties are highly desirable for usage in clinical practice. So too are PRO measures that produce scores that are *actionable*, promoting clinical team decisions to take realistic, efficacious actions on behalf of the patient.[4]

- **It is available in multiple languages.** While not required, PRO measures are made more useful if they are available in multiple languages as they are able to target greater segments of the population. Formal translation procedures should be demonstrated including forward- and backward-translation, and validation to determine semantic and measurement equivalence across language versions.[5]

<u>The utility of PROs</u>

There is accumulating evidence suggesting that routine, formal assessment of PROs in the clinical setting can lead to improved care in many ways.[6] Adding PRO assessment can improve problem identification and patient-physician communication. A systematic review of randomized clinical trials on the impact of feeding back PRO information to clinicians found that in over half of trials care processes were favorably affected.[7] Specifically, PRO feedback increased the number of diagnoses

and notations made in the medical chart and led to more clinician advice, education, and counseling during the patient visit. Effects on patient health status were less frequently assessed and observed. Some studies have also found significant increases in activities to manage the issues identified by patients as problematic and increases in the number of times HRQL is discussed during consultations.[8-11] Less is known about the capacity of PROs to promote patient-centered care (e.g., shared decision making) as few trials have measured such outcomes.[12]

PROs can provide additional information on treatment effects and patient perceptions that are not readily captured by external criteria and clinician-reported outcomes. Given that they represent the patient's subjective feelings and views, it is indeed the patient who is the best source of this information. Furthermore, in some conditions such as cancer, chronic heart failure, and chronic obstructive pulmonary disease, baseline patient-reported well-being can predict objective indicators like survival time.[13-15]

In addition to their use in clinical research and their utility in guiding clinical care, there is increasing interest in using PROs for the purposes of accountability and performance improvement. A PRO can be used to inform derivation of a performance metric, known as a PRO-based performance measure or PRO-PM.[1] A PRO-PM can be based on the scores of a PRO measure alone or in combination with other clinical data (e.g., diagnosis codes). Unlike using PRO data to inform the care of individual patients, data for a PRO-PM are aggregated across an entire health-care entity. For example, a PRO-PM could consist of the percentage of patients in an accountable care organization with an improved depression score as measured by a standardized PRO tool.[1] Using PROs in this way could help address a key challenge of health-care reform – achieving the highest quality care at the lowest possible cost. But in order for this to be realized, two major challenges must first be met.[1] First, PROs have *not* been widely adopted for clinical use outside of research settings; hence, most clinicians, payers, and provider institutions are unfamiliar with them. Second, there is little known about the best set of questions to aggregate for the purpose of measuring performance of a health-care entity. Efforts are currently underway to determine guidelines and best practices for selecting PROs and aggregating PRO data into PRO-PMs. This work is being undertaken by the National Quality Forum (NQF) with funding from the U.S. Department of Health and Human Services.[1,16]

Finally, concerns that measures of the patient experience belong to an inferior category of "soft" outcomes have been superseded by a preponderance of evidence that indicates that the science of PROs meets the same criteria as any other scientific entity, complete with established guidelines and best practices. In establishing this level of credibility for PRO science it has been demonstrated that both laboratory measures (that rely on interpretation[6] or fail to correlate with outcomes patient experience such as death, important morbidity or loss of quality of life) and clinician judgments of the patient experience are less reliable than PROs.[17,18]

Potential sources of measurement error in PROs.

As with any measurement, there are unique sources of error associated with PROs.[2] Patients can misinterpret the meaning of a particular question or its instructions. They may fail to accurately transcribe their answers to the response options or rating scale provided. There are also a number of personal response biases that can jeopardize the accuracy of results. Satisficing is a term used to describe the provision of answers in a manner that is satisfactory, but sub-optimal, involving little to no cognitive effort on the part of the respondent. For example, selecting response options at random, endorsing the status quo, or selecting the same response option for all questions. Patients may also respond to questions in a manner designed to create a certain impression, either favorable (in the case of *socially desirable* responding and "faking good") or unfavorable (in the case of *deviant* responding

and “faking bad”). For instance, a patient may minimize symptom reports to appear healthier (possibly to avoid additional medical tests) or report more severe symptoms to appear sicker (possibly to gain access to more services). These are only a few of the sources of error inherent to PRO measures (a more complete description can be found in Streiner & Norman, 2008).[2]

There are several actions that can be taken to minimize these errors and potential sources of bias, including the following:

- **Screen questions and items.** Use questions and items that appear to be relevant (and non-trivial) to the population of interest.

- **KISS (Keep It Simple & Straightforward).** Use short and easy to understand items and questions with response options that are not overly complex.

- **Use items that have been used before and avoid creating “armchair” items.** Items from previously validated, commonly-used PRO measures are to be preferred over newly-derived items as they are less subject to bias and more likely to have undergone prior testing with patients. While it may be tempting to write new items for a specific purpose (the “armchair approach”), this should be avoided as the clarity and validity of the new items cannot be guaranteed without prior testing with patients. In instances where a new item(s) is required (e.g., no suitable established alternative exists), cognitive testing of the item(s) should occur before widespread administration can be endorsed. Any change in a previously validated item (i.e., wording change, response scale change) will also necessitate cognitive testing with patients as the item is essentially new.

- **Use the information provided.** Using the information that patients provide in a PRO measure will motivate continued responding. Patients who believe that the information they provide will benefit their own clinical care, the clinical care of others, or both will be more vigilant and respond with greater care and accuracy.

- **Triangulate with other measures.** Relying on multiple indicators (both objective and subjective) is most likely to result in the most complete picture of the patient.

<u>Defining a few PRO terms</u>

A glossary of key PRO-related terms is provided in Appendix A. Principal sources of definitions include: the U.S. Food and Drug Administration’s industry guidance on patient-reported outcome measures,[19] Cella and colleagues (2012),[1] Guyatt and colleagues (2002),[20] and Streiner and Norman (2008)[2]

Chapter 2
Patient-Reported Outcome (PRO) Implementation Into Practice

By Michele Y. Halyard, M.D. & Jeff Sloan, Ph.D.

Characteristics of the "perfect" clinical PRO system

For a full discussion of the implementation of PROs into clinical practice, the reader is referred to the reference by Snyder et al which was intended to be a PRO user's guide for those interested in the subject.[21] A review of randomized controlled trials conducted between 1978 and 2007 reviewing the role of PRO data in daily clinical practice suggested a heterogeneity of impact of PROs with an overall conclusion by the authors that the impact was limited. Despite these issues, however, there were some benefits seen amongst the data with the process of care favorably impacted in 65% of studies reviewed and outcome of care improved in 47%.[7] In an integrated, patient-centered practice such as Mayo Clinic, the potential for substantial benefit is real.

Any system designed for collection of PRO data should be based on the practical considerations listed below.[22]

1. What is the purpose of data collection?
Consideration should be given to whether the data is to be used strictly for clinical practice or will there be a desire to mine the data at a later time for research purposes. Consideration should be given to both near- term and long-term data needs. Data collection should be robust, easy, and readily interpretable so as to facilitate the evaluation of its usefulness in clinical practice.

Data may be collected on a one time basis to detect issues that might have otherwise gone undetected. As well, PRO data may be collected at multiple time points to demonstrate changes over time. PRO data collected for individual patient care can also be aggregated across patients and used to evaluate the quality of care within a practice or for comparison of quality of care across providers, as is being done in the United Kingdom (see Chapter 4 for more details on the UK experience).

2. What are the system design considerations?
Design steps for implementing a PRO system should include agreed upon goals for PRO collection, development of system specifications, clinician and patient design feedback, website construction, and usability testing with patients.

3. How will data collection occur?
For many practices, having the patient complete a predetermined questionnaire at the time of the visit in the office is the most convenient. Others prefer to have data collected ahead of arrival to the visit. More literature exists with the collection of PRO data in the outpatient versus inpatient environment, but with the advent of mobile devices, the collection of electronic PRO data is increasingly possible.

Regardless of the practice setting, data collection should be robust, easy, and readily interpretable so as to facilitate the evaluation of its usefulness in clinical practice. Proactive administration of questionnaires can become part of the practice routine and can contribute to increased efficiency.

When patient-reported outcome / quality of life (PRO-QOL) data are to be collected in the clinical practice setting, several factors must be taken into consideration.[23]

- The method of data collection versus available resources is important and will often determine such basics as whether paper or computers will be used. Given the computerized environment of Mayo Clinic, ideally the PRO information should be collected electronically. Not only does this obviate the need to scan in paper later into the record, but it allows the viewing of scores over time such as we now do with laboratory data. As well, electronic capture allows alerting of the clinician to clinically significant changes over time.
- Web-based systems ideally should integrate with an institution's existing electronic medical record (EMR) to facilitate use during the clinical visit. Having the PRO data housed in a separate system from the medical record is a major dissatisfier as discovered in the Mayo Clinic Arizona feasibility pilot study. The clinicians found it difficult to log into and out of a separate system to view PRO data as opposed to having a way to view the data integrated into the EMR.
- Electronic data collection allows patients to enter data remotely at their leisure, i.e. from home, with the potential for electronic alerts (e.g., text messages, email) as reminders.
- Patients must be instructed in how to complete questionnaires. The literature suggests that most people accept electronic data capture as doable and feasible.

One example of a PRO data collection program that is under consideration in the Mayo radiation oncology practice is the web-based Visiontree product. This system is in use at MD Anderson and Cleveland Clinic in Florida. This allows both in-clinic and remote response to PRO questionnaires and also has the capability of displaying which values are of clinical significance. This is also being explored for an ALLIANCE trial for PRO collection being written by Drs. Halyard and Sloan.

4. What questionnaires/tools will be used?

Issues to consider when selecting which PRO instruments to use include:[21]

- whether to use generic or disease-specific questionnaires, profile or preference-based measures, single or multi-item scales, or static (where every patient responds to the same set of questions), or dynamic questionnaires (where questions sets are tailored to the individual).
- need to balance between clinician and patient preferences in selecting the questionnaire and the type of PRO data to collect.
    - PROs can include data on symptom burden, functional limitations, quality of life, health behaviors (e.g., diet, exercise, smoking), and treatment compliance.
    - Physicians may only want to assess issues they know how to treat or manage.
    - For patients, there may be specific symptoms and quality-of-life issues they want to talk to the doctor about.
- Patient burden- A balance must be struck between using questionnaires that capture many PRO items, versus patient burden in filling out forms.

5. Who will review the data and when?

The following represent key issues to consider that will help to formulate the methods by which data are collected:

- Regardless of when the data are collected, it should be reviewed either before or during the patient encounter. Who will review the data and address the issues with the patients must be determined.
- Training of staff must occur on how to interpret the data.
- Changes in workflow should be determined, i.e., when, and by whom, data will be viewed in relationship to the patient visit.
- Determination of what issues will be addressed during the clinical encounter and by whom

The discussion of PRO data with the patient can occur in various ways. Clinicians must be willing to discuss at least some, or if not all, clinically significant issues identified if desired by the patient. Some clinicians will choose to discuss the data with the patient at the time of the encounter. The clinician involved in the discussion may not be limited to the physician overseeing care, but may also include allied health professionals such as mid-levels, nurses, social workers, psychologists, therapists, etc.

6. What will be done with the data? How can scores be interpreted easily and made clinically relevant?

It is imperative that clinicians be able to interpret PRO data easily and that the data be clinically relevant depending on the clinical situation. Training of clinicians on how to interpret scores should be undertaken before implementation in the clinic. Also, obtaining clinician input on the design of the data display in a way that is easily interpretable to the clinician is important . In a qualitative study by Velikova including clinicians using PROs, physicians stated a preference for having cut-off points, grading of responses, and graphs showing trends over time. They also felt that training on interpretation of QOL scores was necessary.[24]

The best method for presenting data to patients has not been thoroughly studied. Issues for consideration include the amount of complexity desired by the patient including whether the patient wants to be presented with just their own individual responses, variability between them and other patients in similar settings, or other descriptions. Brundage et al determined that presentation of QOL information in simple graphs or written texts were preferred over more complex graphical information. Patients were accurate 98% of the time when simple line graphs were used to display data. Furthermore, line graphs were rated highest in both of ease of interpretation and perceived level of helpfulness.[25]

From the actionable standpoint, there are a number of ways that identified PRO issues may be addressed: [21]

- Guidelines can provide information on score meaning - “higher scores mean better functioning” - but such an approach provides no information about the score’s clinical importance or importance to the patient.
- Provision of cut-off scores for “caseness” or levels of severity (e.g., no disability, moderate disability, severe disability), if such data are available for the clinical situation. The usefulness of this approach depends on the sensitivity and specificity of cut-offs, and predictive value will depend on the prevalence of the condition in the population being screened.
- Reference scores from research studies with similar patients, from the general population with the same condition, or from healthy populations can also be helpful by providing a benchmark of other individuals’ scores.[26,27]
    - Such reference scores provide a basis for comparison but do not necessarily indicate what a given score represents for a particular patient

    - Comparison with benchmarks from group data can be problematic because of the significantly larger error of measurement in individuals.
- A more resource intensive approach, but one most likely to be used by Mayo Clinic providers is to have personnel review patient's scores with the patient to clarify and elaborate on problems indicated by the PROs. In our present applications, it became obvious that the interactive review with the patient of the PRO scores garnered more specific and practically actionable information than merely observing the scores.

7. How can patient-reported information be used, including leading to greater clinical action? How can patient-reported measures be used to draw out emotional issues that patients may have surrounding their health condition and its treatment or other issues that may be less frequently addressed in the clinical encounter?

Collection of PRO-QOL data may impact treatment decisions and aid in addressing patient concerns, improve both patient and provider satisfaction, and importantly improve patient QOL.[28] Marshall et al. have identified four areas in which PRO assessment may assist in the clinical practice setting: serving as screening tools, identifying patient preferences to assist clinicians in making informed decisions, improving patient-provider communication, and facilitating shared decision making.[29] As an example, Goodwin et al. have shown that assessment of QOL can be used to guide treatment decisions targeting specific symptoms and psychosocial issues in breast cancer patients.[30] By obtaining data directly from the patients' themselves, it is conceivable that a more robust assessment of patient symptoms can occur. Studies have demonstrated that many patient symptoms go undetected with the usual review of symptoms performed in the clinic, and as well, underestimation of patient symptoms may occur by clinicians.[31-33] An example of the benefit of the use of PRO data comes from Velikova et al. who evaluated the use of QOL data in an outpatient medical oncology setting. More frequent discussion of chronic non-specific symptoms such as sleep, appetite, and fatigue-related problems occurred in the intervention group of the trial where data was seen by the oncology clinicians prior to the patient visit without prolongation of patient encounters. QOL improvement was associated with explicit use of QOL data, including discussion of pain, and role function, with improvement seen in overall and emotional well-being.[10]

Although identification of PRO-related issues is important, action upon those issues is important as well. Marshall et al. performed a literature review looking at the impact of PRO collection on clinical action. Two-thirds of the studies reviewed measured provider behaviors as an outcome. In nine of 17 studies (53%) PRO feedback increased provider detection and management of patient issues in the short term. In the intermediate and long-term time frame, one of two and three of seven studies, respectively, found a positive impact on provider behaviors with PROs seeming to have the most impact in studies of patients with mental health issues. Eight of 10 (80%) of trials showed improvement in provider detection and management.[29]

In the non-oncology setting, PROs are being administered to diabetes patients in the southeast Minnesota Beacon health information technology project. Led by a research team from the Mayo Clinic, the Beacon assessment captures the most critical concerns of the patient and feeds this information forward to health-care providers for use in patient consultations. Potential clinical actions that may help to address these concerns are also provided to clinicians (for more information on the Beacon interface see Chapter 5).

8. Can we use PROs for shared decision making?

When addressing PRO-QOL related issues, the clinician should involve not only the patient, but also the family, caregivers, and other providers when determining the appropriate course of care. It is important to take into consideration the patient's values, preferences, social, or cultural characteristics.[23] Once data are collected, the clinician may choose to share the data with a multi-disciplinary care team to devise a treatment plan for the patient in advance of discussion with the patient. Verhoef et al. described the challenges multi-disciplinary care of rheumatoid arthritis patients, which is similar to the oncology setting, regarding the lack of aligned treatment goals, inadequate focus on daily activities and societal interactions, and lack of significant participation of the patient in setting treatment priorities.[34] PRO-QOL data may be used as part of the discussion of multi-disciplinary team meetings to facilitate communication between clinicians from different backgrounds by providing a structure around which a patient's problems can be discussed and may also lead to improved satisfaction with multidisciplinary care.[34,35]

For examples of Institutions using PRO data collection in clinical practice, please see Appendix B.

Chapter 3
What is Mayo doing / not doing now?

By Victor Montori, M.D., James Naessens, Sc.D., Carrie Thompson, M.D. & Phil Hagen, M.D.

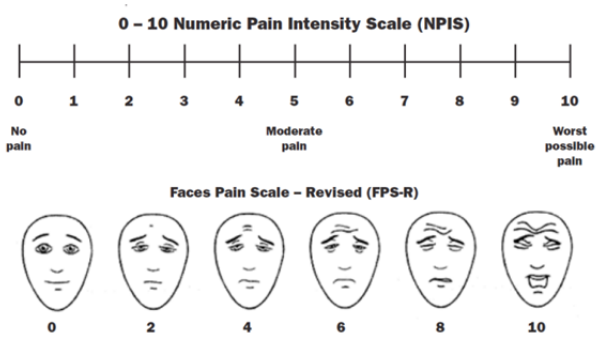
Mayo Clinic has not had an independent institutional policy to give patients voice regarding the experience and outcomes of care. This has historically hindered the depth of analyses available to investigators using our medical records (e.g., Rochester Epidemiology Project). More importantly, it has limited the ability of clinicians to understand the effects that management strategies and the organization of care delivery have on their patients.

Summary of current state

Initiatives that exist are all fairly **embryonic**, and are mostly **motivated by compliance** with accreditation, public reporting, or payer organizations. Either they are in the pilot states of data collection and analysis, or they have robust data collection, but the data are used in a limited way in guiding clinical care, or in analyzing the practice. In preparing this report, it is obvious that **there is no one group that is coordinating** the PRO design, selection, collection, data management, interpretation, and integration into clinical pathways, and there is extreme paucity or limited access to performance or outcome data.

Current experience

Pain – the most visible and widespread effort refers to the routine collection of pain data from patients. Data collection follows a Joint Commission standard (PC.01.02.07) and is inscribed into Mayo Clinic policies (Assessment and Care Planning Across the Continuum of Care - PA.10 http://mayocontent.mayo.edu/ipm/DOCMAN-0000055912, Pain Management - PA.7 http://mayocontent.mayo.edu/ipm/DOCMAN-0000056011).



The care team is responsible for collecting the data. The clinical assistants collect this data as part of a rooming workflow ("Flow Form"), and clinicians should address pain scores in the clinical care and note this in the clinical notes. The data collection instrument is either a verbal request for a 0-10 pain report or the use of a numeric pain intensity scale or faces pain scale tool (see figure – Mayo Clinic form MC3100-03). At Mayo Clinic, The Outpatient Practice Subcommittee and the Accreditation groups are responsible for the routine implementation of pain assessments in the outpatient setting (in the inpatient setting, pain assessment is routinely conducted by nursing, stored in a different location in the medical record and is often part of protocols of care,

particularly in the post-operative period). The clinician is responsible for handling responses to this PRO.

The level of adherence is established by identifying 10 randomly selected patient records per clinical area of interest per quarter. The audit includes completion of pain assessment and plan to address pain scores > 4. The Health Information Management group conducts the audits; the Outpatient Practice Subcommittee reviews the results and responds to noncompliance. For example, as of July 2012, primary care areas assess pain (and address pain > 4) in 70-90% (100%) of their patients, oncology 100% (71%), physical medicine 80% (100%), headache clinic 70% (100%) and pain clinic 100% (100%).

At the time of writing this report, there are no institutional clinical pathways in place to respond to pain scores or to respond to patterns over time. Also, there are no known or published results of the effect of collecting this PRO routinely on the care of patients at Mayo Clinic.

Depression – At Mayo Clinic, the collection of a measure of depression has been part of the implementation of the DIAMOND project, a primary care initiative across Minnesota to improve the care of patients with depression in collaboration with psychiatrists and care managers.[36] The measure selected for assessment and follow-up of depression is the Patient Health Questionnaire or PHQ-9 (a short form, PHQ-2 is used in some areas as a screener).[37] The PHQ-9 asks of symptoms of depression over the last 2 weeks. Different clinical areas use the PHQ-9 with patients following triggers that differ by area.

Adherence to the measure relates to who is responsible for giving the print form to the patient, whether the patient completes the form, whether the form returns to the clinical assistant or clinician, and whether it gets entered into the electronic health record. An implementation in the electronic health record is available, but informal reports suggest that clinicians find it too cumbersome for use at the point of care.

Using the proportion of patients with a diagnosis of depression presenting for a visit at primary care as the denominator, the proportion that receives a PHQ-9 assessment tool varies from 50-90% across sites, and by 15-30% within clinic month-to-month (this is publicly reported by Minnesota HealthScores; top Mayo primary care sites are reporting ~90% adherence with many Minnesota sites reporting 100%). Workflow to follow-up on results of the PHQ-9 vary by clinical area, with the primary care sites in which the DIAMOND project is set up responding with close care coordination and teamwork including a primary care psychiatrist. In other areas, follow-up is not protocolized.

There are improved outcomes from DIAMOND in which PHQ-9 is an important component, but not necessarily only from PHQ-9. In a different practice, introduction of PHQ-9 in psychiatric practices influenced clinical decisions in 93% of more than 6000 patient contacts with treatment changes (change in dose or adding medicines) taking place in more than 40% of these contacts.[38]

Health-related quality of life – Since November 2009, single-item quality of life (QOL) measures are being collected among patients seen in the Division of Hematology (a mix of cancer and non-malignant hematology patients) and entered into the electronic health record. Patients are asked to complete single-item, linear analog self-assessment (LASA) questions to rate their pain, fatigue, and health-related QOL in the past week at the time that they are roomed for their appointment. About a year later, the Department of Medical Oncology followed the same practice. The clinician is responsible for handling responses to this patient-reported outcome (PRO).

As a Quality Improvement Initiative, data from the first three weeks of the project in Hematology were collected and analyzed (data not published, internal use only). There were a total of 1090 assessments. Mean fatigue score was 3.8 (range 0-10, 0=no fatigue, 10=worst you can imagine), with the highest fatigue reported by patients seen by the acute leukemia group and the lowest reported by patients seen in the chronic lymphocytic leukemia group. Pain does not tend to be a significant issue in hematology patients, with a mean score of 2.1 (range 0-10, 0=no pain, 10=worst you can imagine), highest scores reported by patients seen in the non-malignant hematology group and lowest reported by patients seen in the chronic lymphocytic leukemia group. Mean reported QOL score was 7.1 (range 0-10, 0=as bad as can be, 10=as good as can be) with the highest QOL reported by patients seen in the chronic lymphocytic leukemia group and the lowest QOL reported by patients seen in the non-malignant hematology group. These results were presented to the Division of Hematology in March 2010 with suggested "cut-offs" to trigger concern and provide institutional resources for management of pain, fatigue, and common causes of low QOL in malignant hematology patients (e.g., psychosocial concerns, spiritual issues). While a nursing-led comprehensive assessment and intervention triggered by "red flag" PROs is felt to be ideal, this is not feasible at the current time due to staffing and resources.

In Medical Oncology, the data have become highly integrated into the clinical practice with the majority of clinicians indicating they use the data and see value in their practice as a result (unpublished data). Clinicians indicated that they found issues that would have otherwise gone unnoticed by them through the use of these simple questions of fatigue and overall QOL.

As part of the High Value Healthcare Collaborative, there has been an initiative to collect data on overall physical function, mental function, pain and role function on all patients via self-report (the so-called, **"vitals+4"** initiative). The latter measure, role function, resulted from a community-wide assessment in Southeast Minnesota.[39] To our knowledge, this is in the stage of discussion and design, with some urgency as there are initiatives seeking to mandate such measures on a disease-specific fashion.

Orthopedics has been collecting PROs for patients in the Total Joint Registry since at least 2004 for pain and joint function in an irregular fashion (for example, for primary knees from 2008-2010 about 50% had preoperative assessments and ~50% had 6-12 month follow-up assessments, however only ~25% had both). Recently Minnesota Community Measurement (MNCM) selected the Oxford Knee Score and the EuroQOL – 5 Dimensions measure (EQ-5D) as required measures for primary knee replacement patients at baseline and one year after surgery. MNCM plans to extend patient-reported measures to selected spine surgery patients with the Oswestry Disability Index, the Visual Analog Pain Scale and the EQ-5D.

<u>Extent of Patient-Reported Outcome collection in current patient entry forms</u>

Mayo Clinic - Patient Provided Information (PPI)

PPI is a Mayo designed system for gathering information from patients systematically and electronically with a minimum of manual work. This information is necessary for many aspects of the practice – patient care, education, research, and business office functions. It is highly integrated with multiple Mayo Enterprise electronic systems. PPI involves a smart system that recognizes the logging of a patient in the Mayo multiple registration and appointment systems. Based on a rules

engine it determines what forms/information are needed and mails paper forms or electronically "tees up" forms in the patient portal to be completed online. The most commonly used forms include:

- **Adult and Pediatric – Prior Family History (PFH) and Current Visit Information (CVI)** (basic medical information – referring physician, medications, allergies, review of systems, past medical history, family history, functional status, etc.)
- **Gynecology** – specialty specific visit information
- **Campus Wide Authorization** – insurance information, research authorization
- **PHQ – 9** – Depression Screening and Treatment Monitoring
- **AMQ** – Asthma Management Questions
- **ACT** – Asthma Control Test for Disease Management monitoring

Data gathered by PPI systems is largely "elemental" in nature and, thus, computable. Forms can be printed and mailed automatically using smart logic, or on demand at point of care. Mayo has a very large scale distributed network of smart scanners with redundancy ("fail safe") and redundant "virtual servers." The system has been in continuous operation since 1995 and has served more than 1.5 million Mayo patients.

The PPI system is integrated with multiple electronic medical records (i.e. it is EMR agnostic), the data is used in multiple downstream specialty systems, it is used directly in patient care and can be integrated immediately into clinical notes and other clinical systems. PPI has been shown to:

- Enable multiple and innovative care models – face-to-face, telephonic, online such as the Diamond Project
- Save both dollars and FTE because of its efficiency
- Provide elemental data to the Enterprise Data Trust for research

Support compliance with multiple regulatory systems including insurance, research authorization, and meaningful use.

Other patient entry points - There are no other forms of structured entry of PROs in the patient portal or Mayo Clinic Patient App. There are plans to integrate measures of PROs in care plan documents being designed by the Center for Innovation.

Chapter 4
Capturing Overall (Subjective) Patient Functioning

By David T. Eton, Ph.D.

What is overall patient functioning?

In the context of patient-reported outcomes (PROs), we define overall patient functioning in terms of four inter-related concepts – symptoms, functional status, general health perception, and overall quality of life. These *subjective* concepts have been causally linked to more traditional *objective* clinical indicators of health (i.e., physiological, biological, and genetic markers) in conceptual models of perceived health.[40-42] As articulated in these models, symptoms refer to a patient's perception of an abnormal physical, emotional, or cognitive state (e.g., pain, depression, fatigue). Functional status refers to the ability of the patient to perform particular tasks and activities in various domains (e.g., physical, role, social, and psychological). General health is a summarized evaluation of health (physical and mental) that integrates perceptions of symptoms and functional status. Overall quality of life (QOL) is a general appraisal of a patient's subjective well-being or satisfaction with life as a whole. When the QOL evaluation is limited to the context of health and illness it is referred to as health-related quality of life (HRQL).

Why is it important to measure overall patient functioning?

Assessing perceptions of overall health status and patient functioning is important for a number of reasons. At the point-of-care, individual-level patient data can be used to, (1) screen for new health problems, (2) provide regular monitoring of a patient's health status, and (3) enable "patient-centered care" whereby physicians and clinical teams are alerted to the most pressing concerns of each patient, prompting patient-physician discussion of these issues and ultimately informing a treatment plan that best meets the needs of each patient.[12] Furthermore, PRO data aggregated at a group or population level can help determine the comparative effectiveness of different treatments for similar conditions. As indicated in Chapter 1, group-level PRO data can also be used to inform metrics of accountability and to improve performance across service providers. In this manner, patients could compare the outcomes and health-care quality of various providers, enabling more educated decisions about their health care.[12] A few European countries have been using PROs in this manner, most notably the United Kingdom and Sweden. In the UK, providers offering one of four elective interventions (hip, knee, groin/hernia, and varicose vein surgery) are now required by the UK's National Health Service to collect and report PRO data for quality purposes.[16,22] In Sweden, PROs are collected by local health registries with the data used to improve clinical care and for research. The Swedish Board of Health and Welfare requires the use of PROs and makes PRO-based performance data available to its citizens.[16] Both countries' systems make use of disease-specific and generic PRO measures, such as those assessing overall functioning, to measure patient health status. In the United States, one of the more advanced groups using PROs in performance measurement is the Minnesota Community Measurement Program.[16] This group works with state authorities and insurers to develop and endorse valid and reliable PRO-based performance measures (PRO-PMs) for accountability purposes.

Condition-specific (e.g., PHQ-9, Oxford Knee and Hip scores) and generic health status measures (e.g. the EQ-5D) have been endorsed by this program for use in performance metrics.

What are the options for measuring overall patient functioning?

There are a number of brief, easy-to-complete PRO measures available that can be used to obtain a "bottom line" evaluation of a patient's overall subjective health status. A team of patient-reported outcome / QOL (PRO-QOL) scientists from the Department of Health Sciences Research recently reviewed five of the more user-friendly measures available today. The team consisted of Dr. David Eton from the Division of Healthcare Policy and Research, Dr. Kathleen Yost from the Division of Epidemiology, Dr. Jeff Sloan from the Division of Biostatistics, and Dr. Jeanette Ziegenfuss, formerly from the Division of Healthcare Policy and Research and currently at HealthPartners of Minnesota. In the following section, the strengths and weaknesses of the different options are reviewed (see also Table 1), and overall recommendations for assessment are provided. This process of review by a panel of content experts could serve as a model for vetting other PROs being considered for implementation.

A critical examination of options for assessing overall patient functioning

(1) The **Dartmouth COOP Charts** are designed to provide a brief measure of functional status and health-related quality of life (HRQOL) for use in primary-care settings. The COOP Charts consist of 9 charts, each with an illustration and a corresponding question that queries health status over the past 4 weeks (a 2-week recall version is also available). They are intended for use as brief screening measures to be used in clinical practice.[43] Each chart is scored individually. The charts have the advantage of ease, but tend to be less precise and specific than multi-item scales. Accurate duplication of chart illustrations in a web or on-line portal may limit usability. The charts require a user fee (see www.dartmouthcoopproject.org). Recently, Dartmouth has adapted more easily-implemented, single-item assessments featuring a 7-day recall for three issues queried in the Charts (physical health, mental health, and pain). These are being referred to as the "Vitals + 3." These three questions are embedded in other measures as well (see below).

(2) The Medical Outcomes Study **Short Form-12 (SF-12)** is a shortened version of the SF-36 general health status instrument, making it appropriate for rapid assessment of overall patient functioning. Like the SF-36, it was designed for use in health policy evaluations, general population surveys, and clinical research and practice.[44] Some centers and medical departments are currently using the SF-12 to assess patient health at the point of care (see www.dukepersonalizedmedicine.org/patient_care). The SF-12 can be scored to produce physical and mental component summary scores (PCS-12 and MCS-12) using software provided by the instrument developers (QualityMetric Inc.). The software uses data from all 12 items to produce the component scores by applying an algorithm derived from norms of the U.S. general population. The PCS-12 provides an overall summary of physical health status, while the MCS-12 provides an overall summary of mental health status. Population norms are available enhancing the interpretability and utility of its scores. Standard four-week and one-week recall time frames are available for use with version 2 of the SF-12 (see www.qualitymetric.com). The widespread utility of the SF-12, its brevity (2-3 minutes to complete), and the availability of extant normative data make it a reasonable choice to assess patient functioning. However, the SF-12 does require a user fee for the copyrighted version and scoring algorithm (as do all SF instruments administered by QualityMetric). Scoring the measure using anything other than the standard procedures (i.e., "homegrown" scoring) is discouraged owing to the propensity for scoring errors and consequent lack of instrument precision. Instrument and scoring modifications are occasionally made

to SF instruments to improve precision and respondent comprehension. This can result in a lack of equivalence between older and newer versions.[45] Multiple modes of administration including paper, web, portal, and interactive voice response, are available, though certain modes require that data be shared with QualityMetric Inc.

(3) Simple 0 to 10 numerical rating scales can be used to assess overall patient functioning. Numerous studies have utilized these so-called **linear analog self assessments** (or **LASAs**) as an efficient means of assessing patient health status in a variety of domains. Major advantages of the LASA are its convenience, ease of use, and scoring. A single-item LASA can be crafted for virtually any domain of interest. LASAs assessing pain, as well as other specific symptoms, are commonly seen in studies of chronic illness. One potential drawback of the LASA is that verbal descriptors are not used for all points on the rating scale. This may pose a challenge to the interpretation of its scores. LASAs have a history of extensive usage here at the Mayo Clinic, championed by Dr. Jeff Sloan, and are a major part of quality of life evaluations principally within the area of oncology, and to a lesser extent other clinical areas. Given that LASAs are often tailored to particular evaluations, few national normative data are available, with the exception of the pain LASA embedded within the PROMIS Global-10 measure (see below).

(4) A fourth measure for consideration is the EuroQOL group's **EQ-5D** (EuroQOL – 5 Dimensions). The EQ-5D is a widely used generic, preference-based, health status measure. It is brief, requiring users to assess health status in five domains (mobility, self-care, usual activities [e.g., work, study, family], pain/discomfort, and anxiety/depression). In its original form, each functional domain is judged using 3 response categories (no problems, some problems, extreme problems). Based on responses to the five functional domains, respondents are classified into 1 of 243 distinct health states. Respondents also provide a summary rating of their current overall health state on a 0-100 hash-marked visual analogue scale (VAS). To score the EQ-5D, the researcher assigns a single index value to the health state defined by subject responses to the 5 domains. This index value has been previously derived and is anchored to health state preferences of the general population. This value, known as a "utility," summarizes a person's overall health status in a single number ranging from 0 (dead) to 1 (perfect health). A utility score is useful as an outcome measure in clinical studies, for estimating quality-adjusted life years (QALYs) for economic evaluations, and analyses of cost-effectiveness, and for monitoring the health of populations. At the individual, patient-level, the utility score has little clinical meaning. The EQ-5D has been widely tested and used in both general population and patient samples (www.euroqol.org).

The EQ-5D was designed to measure decrements in health. One of its major limitations is that it has been found to suffer from ceiling effects, particularly when used in general population surveys, but even in some patient population settings. Given this, there may be issues with its ability to measure small changes in health, especially in patients with milder conditions. Due to this limitation, the EQ-5D is currently undergoing revision, from its current 3-level response format to a 5-level response format (no, slight, moderate, severe, or extreme problems).[46] Large-scale psychometric testing of this new version (being referred to as the EQ-5D-5L), including comparison to the original (or EQ-5D-3L) is forthcoming. Licensing fees for the use of the EQ-5D measures are determined by the EuroQOL group's Executive Office. A web-version is available, but must link with the EQ-5D's centrally-managed, secure web server. The EuroQOL group does keep a copy of all data collected using the EQ-5D; however, identifying information is stripped to maintain respondent anonymity.

(5) A global patient health assessment is also available from the Patient-Reported Outcomes Measurement Information System (PROMIS). PROMIS is a national resource of publically available

efficient and flexible measurements of symptoms, functioning, and HRQOL, appropriate for a wide variety of chronic diseases and conditions (http://www.nihpromis.org). The **PROMIS Global-10** short form scale consists of 10 items that assess general domains of health and functioning including overall physical health, mental health, social health, pain, fatigue, and overall perceived quality of life. The items can be used to arrive at a "bottom-line" summary indication of health status. Global items like the ones making up the Global-10 have been found to be predictive of health-care utilization and mortality in general and disease-specific, clinical populations.[47] The PROMIS Global-10 includes the most widely-used self-rated health item ("In general, would you say your health is…excellent, very good, good, fair, poor"). A question on overall quality of life ("In general, would you say your quality of life is… excellent, very good, good, fair, poor") is a strong indicator of mental health.[48] The pain item is administered as a 0 to 10 LASA, a measure frequently used in clinical research and practice settings. The other items of the Global-10 have largely been adapted from other frequently used "legacy" measures such as the SF-36; however, some have been modified slightly based on results from extensive qualitative item review by clinicians and patients (i.e., focus groups and cognitive testing).[49] This was done to arrive at a set of items with greater sensitivity and precision than the items as originally worded. The Global-10 items and their response scales appear in Appendix C. Note that the three items in Dartmouth's "Vitals + 3" assessment are embedded within the Global-10 (the global03, global04, and global07 items). The Global-10 is also being pilot tested as a point-of-care measure and value metric at two Harvard-affiliated hospitals (Massachusetts General and Brigham and Women's) in patients treated for coronary artery disease or diabetes.[50]

There is considerable flexibility in scoring the PROMIS Global-10. Each of the individual items can be examined separately to provide specific information about perceptions of physical function, pain, fatigue, emotional distress, social health and general perceptions of health. Furthermore, a recent study from the PROMIS network has supported derivation of two 4-item summary scores: a Global Physical Health (GPH) score and a Global Mental Health (GMH) score.[51] GPH and GMH scores can be easily converted to a T-Score metric (using a look-up table) allowing for comparisons to a general (norm) population. A final useful feature of this measure is the ability to estimate an EQ-5D index score from a linear combination of eight Global-10 items.[52]

The PROMIS Global-10 provides an efficient and flexible assessment of self-reported health and may be useful for large epidemiologic and observational studies for monitoring or assessing the health of populations. A personal communication with one of the PROMIS network's lead investigators, David Cella, Ph.D., has indicated that the use of the Global-10 (or select items within) for the purpose of tracking general health status in a patient population is appropriate. While the PROMIS system of measures has never been used to assess performance of health-care entities, several of its components have been used in this way in the past.[1] Hence, it is conceivable that the Global-10 could inform a PRO-based performance metric, although further testing and validation will likely be required.

The measure takes about 2 minutes to complete. It is in the public domain and can be reproduced and used at no cost. Any electronic data collected can be housed on a local platform and does not require transmission to a third party server. Terms and conditions of use can be found at: www.assessmentcenter.net/documents/PROMIS%20Terms%20and%20Conditions%20v7.3.pdf

PROMIS does not require a signed user agreement.

Table 1: Comparison of assessment measures on select domains

	COOP Charts	SF-12	LASAs	EQ-5D	PROMIS Global-10
Cost to use instrument	-	-	+	-	+
Ease of completion	+	+	+	+	+
National Benchmarks available	?	+	-	+	+
Data does not need to be shared with outside vendor	+	+	+	-	+
Overlap with other centers	+	+	+	+	+
Flexible utility in multiple admin. modes	-	+	+	+	+
Ease of scoring	+	-	+	-	+
Ease of interpreting scores	+	+	-	-	+

"+" = Acceptable, relative advantage; "-" = Drawback, relative disadvantage; "?" = Un-determined

Overall recommendations:

While no studies exist that directly compare the relative performances of the five instruments described, our descriptive analysis does enable us to offer some justifiable recommendations for the assessment of overall patient functioning. We offer both a primary and a secondary recommendation to consider. We believe that either of these options will provide a sound assessment of the target constructs.

Primary recommendation: We recommend the use of the PROMIS Global-10 as a general assessment of overall patient functioning. This decision is supported by the following:

- Appropriate representation of basic domains of functioning: general physical and mental health, social and role activities, overall QOL, pain and fatigue
- Flexible scoring, i.e., responses to single items, aggregate summaries of "global physical health" and "global mental health," T-score conversion to allow comparison with population norms, and estimated index utility score for use in cost-effectiveness analyses
- Fully-anchored response options for most items facilitate interpretability
- Inclusion of the commonly-used, 0-10 pain LASA
- Ease of use: the measure can be completed in about 2 minutes
- Ease of access: the measure exists in the public domain and can therefore be reproduced and used at no cost to its user
- Collected data can be housed on site and requires NO transmittal to another party
- Data collection overlap with other centers (e.g., Dartmouth, Harvard) may facilitate comparisons with other populations
- Implementation of PROMIS measures of overall health into electronic health records has been supported by professional organizations like the Society of Behavioral Medicine and by thought leaders at the National Institutes of Health[4]

Secondary recommendation: The PROMIS Global-10 could also be supplemented with two additional LASA items assessing fatigue and overall QOL. While this would produce a slight overlap with the two items assessing fatigue and QOL on the Global-10 (global02 and global08), this addition would provide added value by enabling comparisons with other Mayo patient samples, especially

within oncology. Currently, many oncology patients at Mayo are being administered LASAs on fatigue, overall QOL, and pain as part of routine clinical evaluation. A considerable amount of archived patient data on these measures in the Mayo environment is also available.

Chapter 5
Next Steps Associated with PRO Roll-Out at Mayo Clinic

By Timothy J. Beebe, Ph.D.

Success Stories at Mayo Clinic

Successful patient-reported outcome (PRO) data collection systems capture PRO data that matter most to patients and feed it forward to clinicians in real time, as care is delivered, and offer results back to patients. Using this definition, examples of successful collection and clinical use of PRO data at Mayo Clinic are few and far between. Nonetheless, a fair amount can be learned from Mayo's experiences collecting pain and depression, as was discussed in Chapter 3. Perhaps the best hope for the ascertainment of best practices within Mayo Clinic is the nascent PRO collection system in oncology care. Whereas the collection of pain and depression was driven largely by outside forces (e.g., Joint Commission, extramural funding), the system in oncology was patient and provider led. Moreover, scant evidence of positive impacts of pain and depression collection in clinical care and outcomes exists. Evidence of the latter in oncology is best offered by the case of a Midwestern farmer, aged 67, who just emerged from intensive treatment – including radiation and chemotherapy - for his prostate cancer. By all classical clinical indicators, his treatment was a success as his imaging was clear, his PSA reports good, and he was back working his farm. No one would have suspected his "stupid thoughts." At one of his follow-up check-ups, a nurse asked him to fill out a newly-implemented, short quality-of-life (QOL) questionnaire. Results showed that his rating had dropped to 5 (on a scale of 10). In cancer patients, a QOL below 5 has been linked to double the risk of death within five years. The farmer's score prompted his oncologist to ask, "What is going on in your life?" That simple question was all the invitation the patient needed to talk about his sleep problems. Every night he thought about death and suicide. As a result of this insight, the patient's doctor convinced him to see a psychiatrist to get the counseling and medication he needed. One month later, at his oncology follow-up appointment, his quality of life score was back up to 8, and life was great.

There may be other examples of this type of impactful PRO collection on the lives of patients and the behaviors of their health-care providers at Mayo. The PRO collection in orthopedics as part of their Total Joint Registry and the work beginning on Vitals+4 initiative as part of the High Value Healthcare collaboration may hold promise. In addition, the PRO-QOL tool that was developed as part of the Beacon community collaborative (see figure on next page) is being piloted in our endocrinology practice and a version of this approach is also being tested as part of a "QOL Integration" concept in Kasson Clinic utilizing the Health Leads (http://www.healthleadsusa.org/) model of integrating social determinants screening into the clinic visit combined with a volunteer staff to help patients connect with resources.

Lessons may be gleaned from these activities as well. The first next step is to conduct a more in-depth analysis of these and other practices within Mayo that are routinely collecting PRO data in order to pull together a listing of lessons learned so that their successes can be diffused as best practices across Mayo.

Success Stories Outside of Mayo Clinic

In addition to more in-depth investigation of Mayo practices, an analogous undertaking should be pursued with our clinical contemporaries who have successfully implemented PRO collection into their practices. A good start to this was offered in a recent peer-reviewed technical report published by The Dartmouth Institute for Health Policy and Clinical Practice, Center for Population Health, titled, "Using Patient-Reported Information to Improve Health Outcomes and Health Care Value: Case Studies from Dartmouth, Karolinska, and Group Health" where the authors conducted in-depth studies of patient-reported measurement systems at these three institutions, largely to demonstrate the clinical utility of such systems, illustrate the feasibility of using PROs in clinical settings, and to ascertain key lessons from these prior activities (http://tdi.dartmouth.edu/images/uploads/tdi_tr_pri_ia_sm.pdf). The lessons learned and the observations made by the authors of the report fell into the broad domains of content, adoption, EHR integration, delivery options, privacy and security, and flexibility and modifiability. Granular examples of each are offered in the report but overall, all three institutions studied have worked to make PRO data collection and the feedback process user-friendly for both patients and clinicians and aim to use PROs to not only offer feedback to improve health, health-care use, and health behaviors, but to improve patient-clinician communication and relations.

Another example of a success story outside of Mayo Clinic is the work recently started by investigators at Partners Healthcare, Harvard Medical School, Brigham and Women's Hospital, and Massachusetts General Hospital where they endeavored to develop a PRO tool for diabetes and CABG patients and collect PRO data from these patients across the continuum of care. They also wanted to focus on the evaluation of data reporting where patient-level PRO results are offered to both patients and clinicians, assess the impact of supplying aggregated data to quality managers on quality improvement activities, and develop an estimation of impact of the new PRO data collection infrastructure on value (quality and cost). Data collection started in March of 2012 and after 6 months of activity, the team observed positive impacts on the patient experience ("doctors should be asking these questions"), the staff experience, particularly the accommodation of PRO collection in their existing clinical work flow, and the physician experience where most agree that there is value in measuring PRO data.

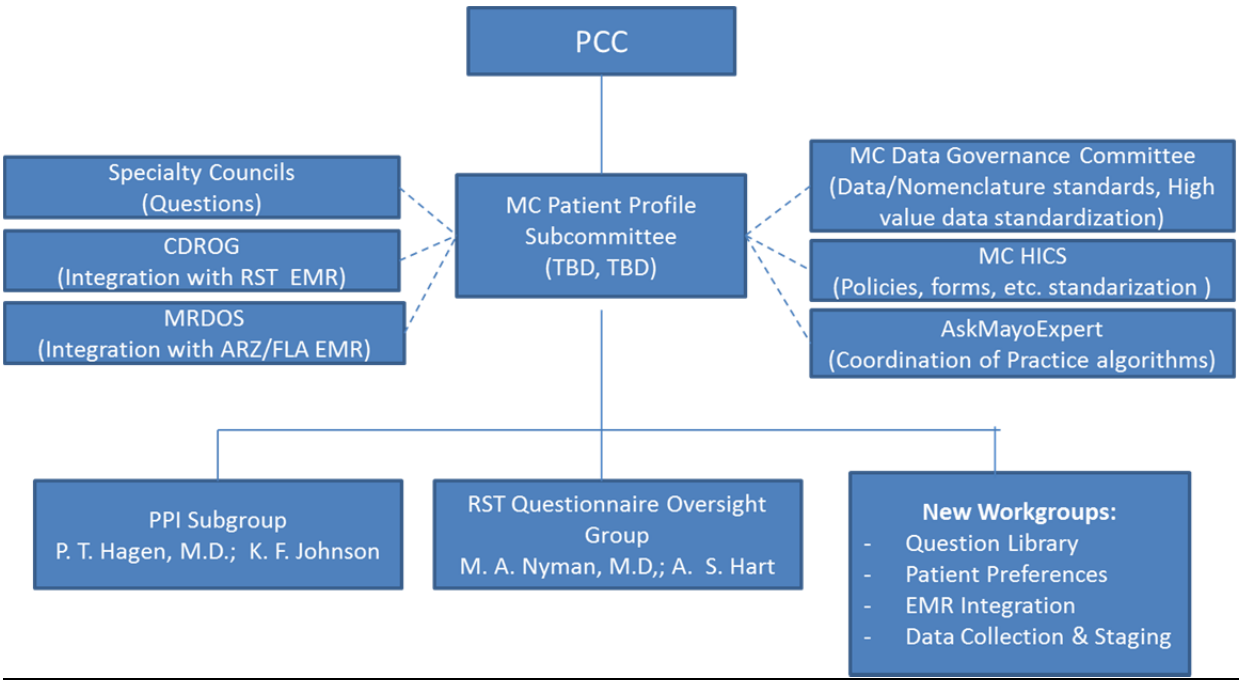
The experiences of these different institutions ought to be more carefully studied, perhaps with a visit to one or more of these sites ourselves. Particular attention might be paid to the experiences of Partners Healthcare as they are similarly driven to harness the power of PRO collection in the quantification of value and use the PROMIS Global-10 in their collection system (as we are recommending). Doing so is important as much has been learned by these various teams and one ought not reinvent the wheel if at all possible.

Engaging Key Stakeholders

It is clear that PRO collection and use is fragmented at Mayo Clinic and little evidence that it has been integrated into electronic medical records or impacted patient care. Efforts to date have suffered from a lack of centralized oversight of the resources utilized to collect PRO data and lack of solid performance metrics. A major commitment must be made across the institution to harmonize the collection of PRO data for the sole purpose of bringing about the highest amount of benefit with the least amount of burden. Fortunately, there are nascent efforts to do just that. The Mayo Clinic Patient Profile Subcommittee, which reports to the Practice Convergence Council, is tasked with defining the strategy for capturing and managing patient-provided information at Mayo, including information collected in electronic medical record clinical forms, risk assessments, and current visit information. It is also charged with creating a road map for the conversion from paper-based, manual forms, to electronic as well as driving the standardization of questions and terminology. The near-term organizational structure of the group is provided below. Clearly, any effort to improve and standardize the collection of PRO data at Mayo Clinic will have to engage this group as a key stakeholder as many efforts in this space are currently underway.

Any PRO data collection system has to be user-centric. As such, every effort to engage both patients and practicing clinicians ought to be made. Either formal focus groups or semi-structured discussion groups (or both) will be held with providers and patients in multiple settings, although the actual number is not yet known. The goal of these sessions will be to land upon a core set of PRO measures that address areas most important to patients and providers, but not disrupt clinical flow. Feedback from this formative work will be compiled by a core team of experts in psychometrics, survey design, quality of life measurement, health literacy, and behavioral psychology. In addition, an oversight group that will be charged with reviewing the work of the core team and will include specialists, primary care providers, public health staff, and HIT experts. To date, there is no formal mechanism for prioritizing or vetting questions at Mayo. Clearly, a team is needed to perform this task. The Center for the Science of Health Care Delivery will work closely with the Practice Convergence Council to pull together such a team. Whether the Center integrates within the existing structure evinced by the diagram below or the converse is unclear at this point and should be discussed forthwith.

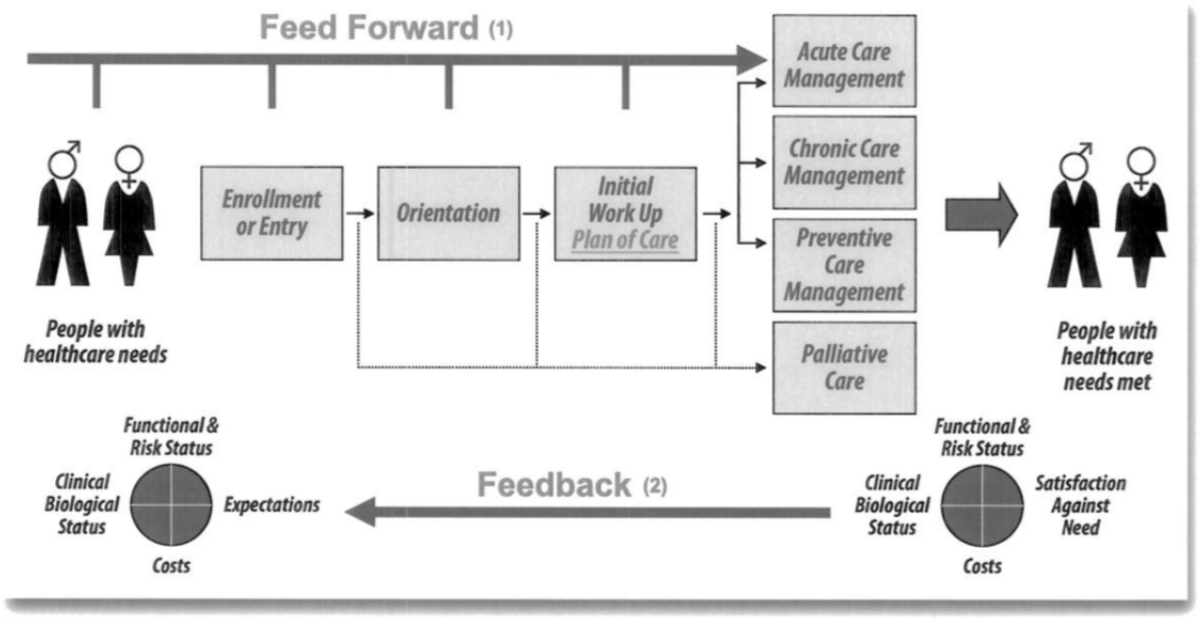
The results of the work will be presented to various specialty councils, the Mayo Clinic Clinical Practice Committee, and to other internal stakeholders based on the direction of the Patient Profile Subcommittee.



The Future of PRO Data Collection at Mayo Clinic

As was stated at the beginning of this chapter, successful PRO data collection systems capture PRO data that matter most to patients and feed it forward to clinicians in real time, as care is delivered, and offer results back to patients. The figure below illustrates these ideals, along with how PRO data collection can be combined with various other data streams (e.g., diagnostic tests, clinical notes, and billing data) to contribute not only to patient care, but to program improvement and research. An effective PRO data collection system should also take into account the "whole" person including the social and behavioral influences on their health. The Beacon PRO-QOL tool we developed collects and informs providers on how to act upon these types of PRO data (see figure on page 25). It is to these ideals that Mayo Clinic ought to strive.

FIGURE 1



Feed Forward and Feedback Data Flow

Diagram illustrating use of feed forward and feedback data in flow of care for patient care, improvement and research.
1. Feed forward to keep data about the patient with the patient as care is delivered (in multiple settings) over time.
2. Feedback to provide summary data on clinical populations to improve care in individual programs, collaborative networks and to provide research data base.

Copied from Nelson et al. (2012)[53]

All three institutions studied in the aforementioned Dartmouth report on PRO collection – Dartmouth, Karolinska, and Group Health – have moved "upstream" in their data collection by encouraging patients to enter data from their own homes using secure web services rather than having them offer PRO data solely in the office setting. However, Mayo Clinic should not stop short of getting patients to input PROs anywhere (patient app, email, SMS, kiosks throughout the clinic). If PRO collection is limited only to our Patient Provided Information forms (PFH, CVI), we run the risk of only having PROs when patients are coming for appointments which may end up being during exacerbations of chronic conditions or when acute conditions appear. We will not capture improvement and therefore value. As such, PRO collection has to be ubiquitous.

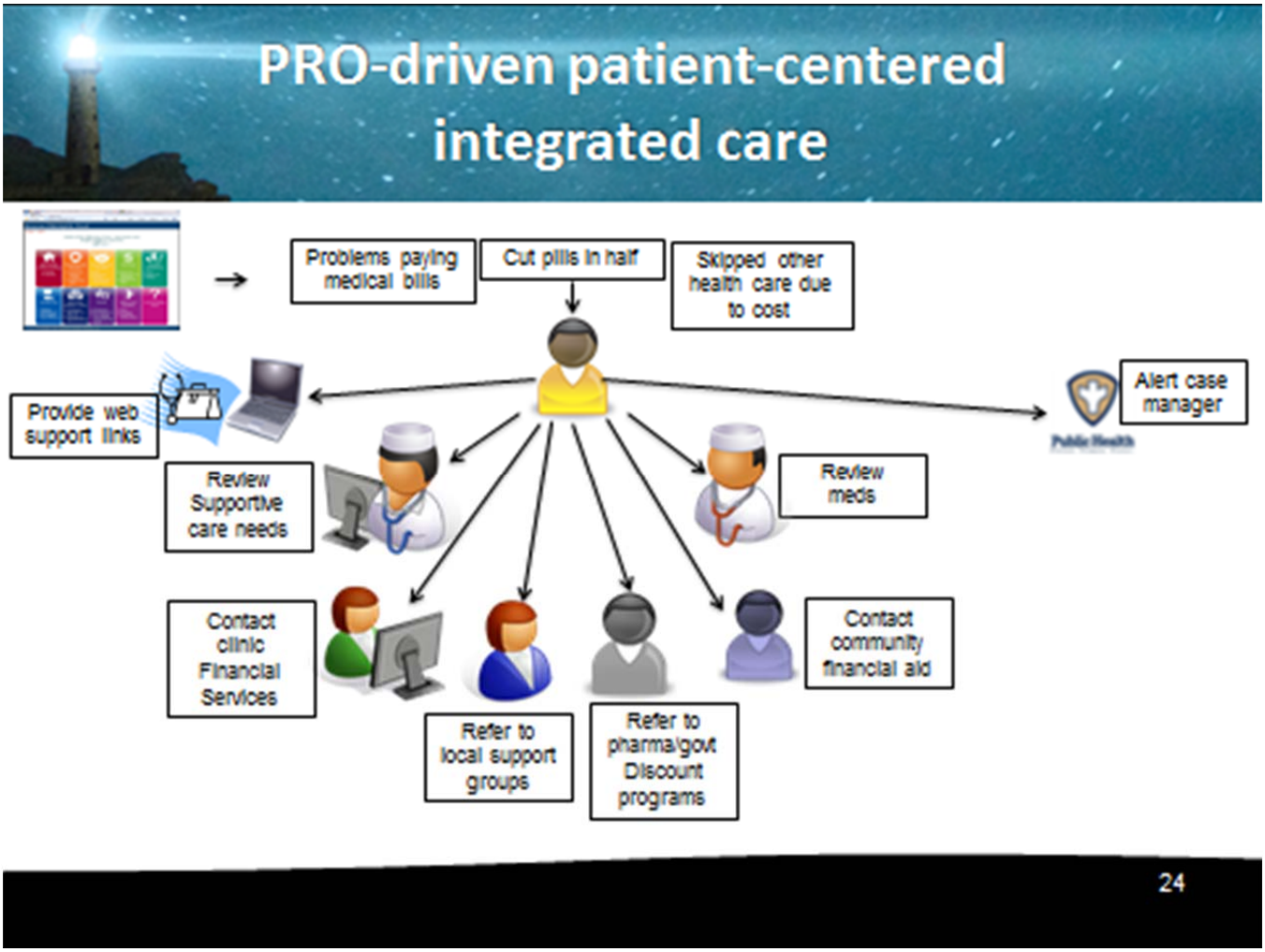
While acknowledging the above, there are some operational and attitudinal constraints with the use of some data collection platforms. In the table below, Eyal Zimlichman, MD, MSc, from the Division of General Internal Medicine, Brigham and Women's Hospital, and the Department of Clinical Affairs, Partners Healthcare, Harvard Medical School, lists the pros and cons of the various methods of collecting PRO data in the clinical setting. While electronic collection of PRO data via computer or IVR (a system which allows a computer to detect voice and/or keypad inputs via telephone) brings about a myriad of potential virtues such as convenience and clinical feasibility as well as instantaneous scoring and the full use of computer-adaptive testing (CAT) - a particular area of focus within the PROMIS initiative - it is clear from the table that there is no one electronic mode of collection that emerges as optimal.

Application	Pro	Con
Patient Portals	• Patient fills out from home • Patient can fill out when convenient • Simple data integration • Promotes adoption of Patient Gateway	• Only 15% adoption rate
Interactive voice response (IVR)	• Widely used • Can have Patient Call-back option • Cheap solution for large populations • Simple data integration	• May be off-putting to patients • Limitations to voice recognition capabilities
Human Operator Phone Calls	• Widely used • Human response	• Acquiescence bias • More expensive than IVR • Cannot feed data back to provider in real-time
Email	• Patient can fill out when convenient • Simple data integration	• Limited to tech literate patients
Mobile Phone Apps	• Patient can fill out from anywhere • Simple data integration	• Limited to tech literate patients
Kiosks/Tablets	• Patient fills out while waiting for appointment • Simple data integration • Can be leveraged for other purposes	• Collects Data at single point of care • Potential for provider to introduce a selection bias
Paper	• Tried and tested method	• Want to move into electronic data gathering • Cannot control (or keep track of) the point of time at which patient answers • Traditionally turns up low compliance rate • Cannot feed data back to provider in real-time

While secular trends are constantly changing with regard to patient and provider preferences for various electronic and paper methods of data collection, it is the paper form that will likely remain the collection method of choice for a number of reasons. First, a significant subset of health-care systems lack the resources and infrastructure to routinely collect and utilize computerized PROs and small, financially-challenged physician groups and rural hospitals and clinics lag behind their larger, better financed counterparts in the adoption of health information technology that would support electronic collection of PROs. Until this gap in resources and infrastructure is closed, paper-based alternatives must be made available to small or financially-challenged clinical practices (e.g., rural and inner city clinics). Second, paper-based collection of PROs does not require computer or telephone access and are portable, allowing patients to return to the task easily after interruptions. This is why paper-based questionnaires remain the standard in general and clinical population sample surveys.

The equivocal nature of the research findings on the relative merits of electronic- versus paper-based collection of PRO information suggests that more attention ought to be paid to the content of the PRO questionnaire rather than the mode of administration or delivery platforms. With both activities, every effort must be made to include both patients and providers in the design and testing phases of the data collection tool, the method of collection, and the venues for reporting and general dissemination. It is standard practice in survey research and instrument design to include subjects in the design of the questionnaire and data collection protocol. Formal piloting and pretesting can take various forms. For example, candidate PRO measures can be subjected to a series of cognitive interviews with a sample of patients where they react to various question forms and/or methods of collection, comparing differences in meanings, burden, willingness, and ability to provide the information necessary to give an adequate answer to the question. Data collected from these cognitive interviews can provide a unique understanding of the cognitive processes involved in responding to the candidate PRO measures collected via different modes and patient receptivity to them. One can also field test the PRO instrument and data collection methods with a small sample of patients from the community

where each participant receives a “questionnaire about the questionnaire” asks respondents to rate the length and difficulty of the interview schedule as well as identify any problematic items. In the end and as a PRO data collection system is deployed, it may be best to have multiple methods of collection and let the patient choose the method that comports with his or her own specific preferences by asking, “what method works best for you?” or “what method would you prefer” as a matter of course.



Chapter 6
How Do We Influence Others
(i.e., Minnesota Community Measures, Center for Medicare & Medicaid Services, talking to stakeholders such as insurers and the state)?

By Victor Montori, M.D., James Naessens, Sc.D, and Douglas Wood, M.D.

Specific health-care reform activities and efforts to increase the value of health care have called for the use of more "patient-centric" measures. In their Health Affairs article in May, 2010, McClellan and colleagues identify patient-reported outcomes (PROs) as a key measure of care effectiveness for evaluating accountable care organizations.[54] Once Mayo Clinic has developed an independent institutional policy to give patients voice regarding the experience and outcomes of care and establishes an internal group to help coordinate activities, we should take active measures to encourage the use of PROs in health-care delivery, and influence the selection of specific measures. In its focus on "the needs of the patient," Mayo Clinic should continue to advocate that PROs should form the basis for determining health-care value. Potential groups we want to influence include other providers, governmental and other policy-making groups and patients and payers:

a. Other providers
    i. High Value Healthcare Collaborative
    ii. Medical Group Management Association
    iii. Hospital Groups (University Healthsystem Consortium, Minnesota Hospital Association, American Hospital Association)
    iv. Physician Groups (Minnesota Medical Association, Specialty societies)
b. Governmental/quasi-governmental bodies
    i. Center for Medicare and Medicaid Services and federal agencies (Department of Health and Human Services, Agency of Healthcare Research & Quality, Center for Disease Control and Prevention)
    ii. Patient-Centered Outcomes Research Institute (PCORI)
    iii. Minnesota Community Measurement
    iv. Minnesota Department of Health
    v. Minnesota Department of Human Services
    vi. Minnesota Legislature (especially relevant committees); other state legislatures may also be relevant in the future (neither Arizona or Florida appear ready, but Wisconsin is close)
    vii. Other policy bodies (JCAHO, National Quality Forum, National Committee for Quality Assurance)
c. Other stakeholders
    i. Payers (including large, self-insured businesses and large plans)
    ii. Patient advocates (e.g., AARP)
    iii. Buying groups (Minnesota Health Action Group in Minnesota, Business Group on Health)

There are multiple avenues to have influence on the use of patient-provided outcomes in health-care delivery and the assessment of value. Mayo should develop a communications plan to coordinate and clarify the message that PROs are important and should be the principal assessment in determining value, but the analysis and interpretation of PROs can be complex and need to be performed in a scientific fashion. This communications plan should include the following:

a. White paper
b. Publications
c. Presentations
d. Response to regulations
e. Specialty workgroups
f. Popular press (gray literature) articles and OpEd pieces

However, it is important to emphasize the proviso at the top of this section. This should occur after Mayo Clinic has developed a robust enough system to collect, use, and analyze PROs. Otherwise, the expertise with use, cost and organization of implementation aspects, evidence of effectiveness and testimonials of usefulness from patients and clinicians will not be there or seem contrived, undermining the efforts to communicate the importance and value of PROs with integrity and credibility. Given the urgency to achieve the latter, it is critical to move forward swiftly with making the Mayo Clinic practices the most impacted by the patient voice in the world.

Mayo Clinic, in its participation in the High Value Healthcare (HVHC) Collaborative, can extend the use of PROs internally and at the same time help to influence the adoption of PROs in clinical practice elsewhere. Further development of payment policy and measurement strategies that rely on PROs as the most patient-centric measures of health would be facilitated by our HVHC collaboration on a national level. Partnerships with the Institute for Clinical Systems Improvement and Minnesota Community Measurement Project would be helpful to Mayo in efforts to adopt more consistent measurement strategies and operations in Minnesota. The advantage to Mayo in reducing duplication of effort in gathering and reporting PRO data, or more concerning, creating parallel measurement strategies and requirements, is a significant reason for Mayo to take the lead in trying to influence others in adopting PRO measures.

Mayo should rapidly develop a consistent clinical PRO measurement strategy that can be included in all public communications about our commitment to patient-centered care and value. These efforts should be aimed at policy makers, payers and provider groups, including the Minnesota Medical Association and other specialty societies, Minnesota Hospital Association and Minnesota Health Plans. Mayo should also work to influence the development of public policy in both legislative and regulatory arenas. This would include presentation to legislative committees and state departments. As the Minnesota Health Insurance exchange is created, there is an opportunity for Mayo to influence inclusion of PROs in measurement of provider and plan performance.

References

1. Cella D, Hahn EA, Jensen SE, Butt Z, Nowinski CJ, Rothrock N. *Methodological issues in the selection, administration, and use of patient-reported outcomes in performance measurement in health care settings: Commissioned paper #1*. Washington, DC: National Quality Forum, 2012.

2. Streiner DL, Norman GR. *Health measurement scales: A practical guide to their development and use*. 4th ed. Oxford: Oxford University Press; 2008.

3. Vickers AJ. Multiple assessment in quality of life trials: how many questionnaires? How often should they be given? *J Soc Integr Oncol.* Summer 2006;4(3):135-138.

4. Glasgow RE, Kaplan RM, Ockene JK, Fisher EB, Emmons KM. Patient-reported measures of psychosocial issues and health behavior should be added to electronic health records. *Health Aff (Millwood).* Mar 2012;31(3):497-504.

5. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Eval Health Prof.* Jun 2005;28(2):212-232.

6. Hahn EA, Cella D, Chassany O, Fairclough DL, Wong GY, Hays RD. Precision of health-related quality-of-life data compared with other clinical measures. *Mayo Clin Proc.* Oct 2007;82(10):1244-1254.

7. Valderas JM, Kotzeva A, Espallargues M, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res.* Mar 2008;17(2):179-193.

8. Detmar SB, Muller MJ, Schornagel JH, Wever LD, Aaronson NK. Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial. *Jama.* Dec 18 2002;288(23):3027-3034.

9. Mazonson PD, Mathias SD, Fifer SK, Buesching DP, Malek P, Patrick DL. The mental health patient profile: does it change primary care physicians' practice patterns? *J Am Board Fam Pract.* Sep-Oct 1996;9(5):336-345.

10. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol.* Feb 15 2004;22(4):714-724.

11. Wasson J, Keller A, Rubenstein L, Hays R, Nelson E, Johnson D. Benefits and obstacles of health status assessment in ambulatory settings. The clinician's point of view. The Dartmouth Primary Care COOP Project. *Med Care.* May 1992;30(5 Suppl):MS42-49.

12. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? *Qual Life Res.* Feb 2009;18(1):115-123.

13. Eton DT, Fairclough DL, Cella D, Yount SE, Bonomi P, Johnson DH. Early change in patient-reported health during lung cancer chemotherapy predicts clinical outcomes beyond those predicted by baseline report: results from Eastern Cooperative Oncology Group Study 5592. *J Clin Oncol.* Apr 15 2003;21(8):1536-1543.

14. Fan VS, Curtis JR, Tu SP, McDonell MB, Fihn SD. Using quality of life to predict hospitalization and mortality in patients with obstructive lung diseases. *Chest.* Aug 2002;122(2):429-436.

15. Konstam V, Salem D, Pouleur H, et al. Baseline quality of life as a predictor of mortality and hospitalization in 5,025 patients with congestive heart failure. SOLVD Investigations. Studies of Left Ventricular Dysfunction Investigators. *Am J Cardiol.* Oct 15 1996;78(8):890-895.

16. Deutsch A, Smith L, Gage B, Kelleher C, Garfinkel D. *Patient-reported outcomes in performance measurement: Commissioned paper on PRO-based performance measures for healthcare accountable entities.* Washington, DC: National Quality Forum, 2012.

17. Fromme EK, Eilers KM, Mori M, Hsieh YC, Beer TM. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. *J Clin Oncol.* Sep 1 2004;22(17):3485-3490.

18. Justice AC, Rabeneck L, Hays RD, Wu AW, Bozzette SA. Sensitivity, specificity, reliability, and clinical validity of provider-reported symptoms: a comparison with self-reported symptoms. Outcomes Committee of the AIDS Clinical Trials Group. *J Acquir Immune Defic Syndr.* Jun 1 1999;21(2):126-133.

19. Center for Drug Evaluation and Research. *Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. Rockville, MD: US Food and Drug Administration; 2009.

20. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc.* Apr 2002;77(4):371-383.

21. Snyder CF, Aaronson NK. Use of patient-reported outcomes in clinical practice. *Lancet.* Aug 1 2009;374(9687):369-370.

22. Halyard MY. The use of real-time patient-reported outcomes and quality-of-life data in oncology clinical practice. *Expert Rev Pharmacoeconomics Outcomes Res.* 2011;11:561-570.

23. Lohr KN, Zebrack BJ. Using patient-reported outcomes in clinical practice: challenges and opportunities. *Qual Life Res.* Feb 2009;18(1):99-107.

24. Velikova G, Awad N, Coles-Gale R, Wright EP, Brown JM, Selby PJ. The clinical value of quality of life assessment in oncology practice-a qualitative study of patient and physician views. *Psychooncology.* Jul 2008;17(7):690-698.

25. Brundage M, Feldman-Stewart D, Leis A, et al. Communicating quality of life information to cancer patients: a study of six presentation formats. *J Clin Oncol.* Oct 1 2005;23(28):6949-6956.

26. Gutteling JJ, Darlington AS, Janssen HL, Duivenvoorden HJ, Busschbach JJ, de Man RA. Effectiveness of health-related quality-of-life measurement in clinical practice: a prospective, randomized controlled trial in patients with chronic liver disease and their physicians. *Qual Life Res.* Mar 2008;17(2):195-205.

27. Santana MJ, Feeny D, Johnson JA, et al. Assessing the use of health-related quality of life measures in the routine clinical care of lung-transplant patients. *Qual Life Res.* Apr 2010;19(3):371-379.

28. Hilarius DL, Kloeg PH, Gundy CM, Aaronson NK. Use of health-related quality-of-life assessments in daily clinical oncology nursing practice: a community hospital-based intervention study. *Cancer.* Aug 1 2008;113(3):628-637.

29. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract.* Oct 2006;12(5):559-568.

30. Goodwin PJ, Black JT, Bordeleau LJ, Ganz PA. Health-related quality-of-life measurement in randomized clinical trials in breast cancer--taking stock. *J Natl Cancer Inst.* Feb 19 2003;95(4):263-281.

31. Basch E. The missing voice of patients in drug-safety reporting. *N Engl J Med.* Mar 11 2010;362(10):865-869.

32. Basch E, Abernethy AP. Supporting clinical practice decisions with real-time patient-reported outcomes. *J Clin Oncol.* Mar 10 2011;29(8):954-956.

33. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. *Am J Manag Care.* Aug 2008;14(8):530-539.

34. Verhoef J, Toussaint PJ, Zwetsloot-Schonk JH, Breedveld FC, Putter H, Vliet Vlieland TP. Effectiveness of the introduction of an International Classification of Functioning, Disability and Health-based rehabilitation tool in multidisciplinary team care in patients with rheumatoid arthritis. *Arthritis Rheum.* Mar 15 2007;57(2):240-248.

35. Callahan MB. Using quality of life measurement to enhance interdisciplinary collaboration. *Adv Ren Replace Ther.* Apr 2001;8(2):148-151.

36. Lauren Crain A, Solberg LI, Unutzer J, et al. Designing and Implementing Research on a Statewide Quality Improvement Initiative: The DIAMOND Study and Initiative. *Med Care.* Feb 8 2012.

37. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* Sep 2001;16(9):606-613.

38. Duffy FF, Chung H, Trivedi M, Rae DS, Regier DA, Katzelnick DJ. Systematic use of patient-rated depression severity monitoring: is it helpful and feasible in clinical psychiatry? *Psychiatr Serv.* Oct 2008;59(10):1148-1154.

39. Severson MA, Wood DL, Chastain CN, et al. Health reform: a community experience using design research as a guide. *Mayo Clin Proc.* Oct 2011;86(10):973-980.

40. Ferrans CE, Zerwic JJ, Wilbur JE, Larson JL. Conceptual model of health-related quality of life. *J Nurs Scholarsh.* 2005;37(4):336-342.

41. Sprangers MA, Sloan JA, Barsevick A, et al. Scientific imperatives, clinical implications, and theoretical underpinnings for the investigation of the relationship between genetic variables and patient-reported quality-of-life outcomes. *Qual Life Res.* Dec 2010;19(10):1395-1403.

42. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *Jama.* Jan 4 1995;273(1):59-65.

43. Nelson EC, Wasson JH, Johnson DJ, Hays RD. Dartmouth COOP functional health assessment charts: Brief measures for clinical practice. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. New York: Lippincott-Raven; 1996:161-168.

44. Ware JH, Kosinski M, Turner-Bowker DM, Gandek B. *How to score version 2 of the SF-12 health survey*. Lincoln, RI: Quality Metric Inc.; 2002.

45. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. *Ann Med.* Jul 2001;33(5):350-357.

46. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* Dec 2011;20(10):1727-1736.

47. Bjorner JB, Fayers P, Idler E. Self-rated health. In: Fayers P, Hays R, eds. *Assessing Quality of Life in Clinical Trials*. 2nd ed. New York: Oxford Press; 2005:309-323.

48. Lorenz KA, Cunningham WE, Spritzer KL, Hays RD. Changes in symptoms and health-related quality of life in a nationally representative sample of adults in treatment for HIV. *Qual Life Res.* Aug 2006;15(6):951-958.

49. DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care.* May 2007;45(5 Suppl 1):S12-21.

50. Zimlichman E. *Patient Reported Outcomes: Measuring patient-centered healthcare value across the continuum of care at Partners Healthcare System.* Paper presented at the International Society for Quality in Health Care. Geneva, Switzerland, October 22, 2012.

51. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Qual Life Res.* Sep 2009;18(7):873-880.

52. Revicki DA, Kawata AK, Harnam N, Chen WH, Hays RD, Cella D. Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample. *Qual Life Res.* Aug 2009;18(6):783-791.

53. Nelson EC, Hvitfeldt H, Reid R, et al. *Using Patient-Reported Information to Improve Health Outcomes and Health Care Value: Case Studies from Dartmouth, Karolinska, and Group Health: Technical Report.* June, 2012: Dartmouth Institute for Health Policy and Clinical Practice.

54. McClellan M, McKethan AN, Lewis JL, Roski J, Fisher ES. A national strategy to put accountable care into practice. *Health Aff (Millwood).* May 2010;29(5):982-990.

55. Hagen PT, Turner D, Daniels L, Joyce D, Medical Records Institute. *Very large-scale distributed scanning solution for automated entry of patient information*. Newton, MA: MRI; 1998.

56. St Sauver JL, Hagen PT, Cha SS, et al. Agreement between patient reports of cardiovascular disease and patient medical records. *Mayo Clin Proc.* Feb 2005;80(2):203-210.

57. Lim LS, Williams DE, Hagen PT. Validation of a five-point self-rated stress score. *Am J Health Promot.* Jul-Aug 2005;19(6):438-441.

58. Snyder CF, Jensen R, Courtin SO, Wu AW. PatientViewpoint: a website for patient-reported outcomes assessment. *Qual Life Res.* Sep 2009;18(7):793-800.

59. Abernethy AP, Herndon JE, 2nd, Wheeler JL, et al. Feasibility and acceptability to patients of a longitudinal system for evaluating cancer-related symptoms and quality of life: pilot study of an e/Tablet data-collection system in academic oncology. *J Pain Symptom Manage.* Jun 2009;37(6):1027-1038.

60. Berry DL, Blumenstein BA, Halpenny B, et al. Enhancing patient-provider communication with the electronic self-report assessment for cancer: a randomized trial. *J Clin Oncol.* Mar 10 2011;29(8):1029-1035.

61. Wolpin S, Berry D, Austin-Seymour M, et al. Acceptability of an Electronic Self-Report Assessment Program for patients with cancer. *Comput Inform Nurs.* Nov-Dec 2008;26(6):332-338.

62. Basch E, Artz D, Iasonos A, et al. Evaluation of an online platform for cancer patient self-reporting of chemotherapy toxicities. *J Am Med Inform Assoc.* May-Jun 2007;14(3):264-268.

63. Vickers AJ, Savage CJ, Shouery M, Eastham JA, Scardino PT, Basch EM. Validation study of a web-based assessment of functional recovery after radical prostatectomy. *Health Qual Life Outcomes.* 2010;8:82.

64. Kearney N, McCann L, Norrie J, et al. Evaluation of a mobile phone-based, advanced symptom management system (ASyMS) in the management of chemotherapy-related toxicity. *Support Care Cancer.* Apr 2009;17(4):437-444.

65. Gurland B, Alves-Ferreira PC, Sobol T, Kiran RP. Using technology to improve data capture and integration of patient-reported outcomes into clinical care: pilot results in a busy colorectal unit. *Dis Colon Rectum.* Aug 2010;53(8):1168-1175.

Acknowledgements

We are grateful to the following individuals for the feedback and comments that they provided on an earlier draft at three stakeholder meetings held on November 2nd, 6th, and 13th, 2012 at Mayo Clinic: Mark Nyman, M.D. (General Internal Medicine & chair of the Questionnaire Oversight Workgroup), Philip Hagen, M.D. (Preventive Medicine & chair of the Patient Provided Information Subgroup), Scott Okuno, M.D. (Oncology & chair of the Outpatient Practice Subcommittee of the Clinical Practice Committee), Stephen Sponsel (Director of Mayo Media Support Services), Julie Prigge (Director of Mayo Patient e-Health Services), and Robert Dimler (Mayo patient advocate). We also thank Dr. Lila Finney Rutten, Associate Scientific Director for the Population Health Science Program, and Ms. Kathleen Harrington, Division chair of Government Relations for their thoughtful comments on an earlier draft. Finally, we appreciate the administrative assistance of Mr. John Smith, Ms. Lacey Hart, Ms. Amber Kahnke, and Ms. Diane Olson from the Center for the Science of Health Care Delivery and Ms. Sara Hobbs-Kohrt from the Division of Health Care Policy & Research for assistance on formatting the report.

Appendix A

Glossary of Patient-Reported Outcome Terms

Clinical significant difference: A difference in score on a PRO measure that is perceived by patients to be noticeably beneficial or harmful, and which would lead a clinician to consider a change in treatment or care. The smallest change in score that can be regarded as important is referred to as the minimally important difference (MID).

Cognitive testing: Qualitative research tools designed to determine whether concepts and items are understood by patients in the same way that the instrument developer intends.

Concept: The specific measurement goal (the "thing" that is to be measured by a PRO instrument)

Construct: An underlying latent factor that helps explain the relationships among a set of observations, behaviors, or attributes. Constructs are higher level abstractions that cannot be directly observed, but can be operationalized through the measurement of constituent indicators. For example, "anxiety" is a construct; pacing, sweaty palms, difficulty concentrating, and tachycardia are all indicators of the construct of anxiety.

Domain: A sub-concept represented by a score of an instrument that measures a larger concept that is comprised of multiple domains.

Health-related quality of life: A multi-faceted concept that represents the patient's general perception of the effect of illness and treatment on physical, psychological, and social aspects of life.

Item: An individual question, statement, or task (and its standardized response options) that is evaluated by the patient to address a particular concept.

Linear Analog Self-Assessment (LASA): A simple rating scale designed to elicit a direct quantitative estimate of the magnitude of a single concept or attribute using a 0 to 10 numerical rating scale with descriptive verbal anchors at each extreme. Some LASAs also make use of an intermediate verbal anchor. Respondents are required to circle the number corresponding to their perceived state.

Patient-reported outcome (PRO): Any report of the status of a patient's health condition, health behavior, or experience with health care that comes *directly from the patient* without amendment or interpretation by a clinician or anyone else.

Patient-reported outcome-based performance measure (PRO-PM): A performance measure that is based on patient-reported outcome data aggregated for an accountable health-care entity. PRO-PMs can be based on PRO data alone or in combination with other clinical indicators (e.g., diagnosis codes).

Proxy-reported outcome: A measurement based on a report by someone other than the patient reporting as if he or she is the patient. A proxy-reported outcome is not a PRO and is less accurate at assessing internal states that can only be known by the patient.

Questionnaire: A set of questions or items shown to a respondent to get answers for research or clinical purposes. Terms sometimes used synonymously include, *instrument*, *survey*, *tool*, and *measure*.

Rating scale: The system of numbers or verbal anchors by which a value or score is derived for an item.

Recall period: The period of time patients are asked to consider in responding to a PRO item or question.

Score: A number derived from a patient's response to items on a questionnaire. Scores can be computed for individual items, domains, or concepts, or as a summary of items, domains, or concepts.

Visual Analog Scale (VAS): A simple rating scale designed to elicit a direct quantitative estimate of the magnitude of a single concept or attribute using a line of fixed length (usually 100 millimeters) with descriptive verbal anchors at each extreme and no words describing intermediate positions. Respondents are required to place a mark on the line corresponding to their perceived state.

Appendix B

Centers Developing or Utilizing Electronic Patient-Reported Outcome Data Collection in the Clinic

Institution	Reference	Using in practice Y/N	System	Findings
Mayo Clinic	Hagen et al.[55] St. Sauver et al.[56] Lim et al.[57] Personal Communication	Yes – In use continually since 1995. Multiple questionnaires - >1.5 million patients	Mayo Designed – Integrates with G.E. and Cerner EMRs (EMR agnostic). Integrated with multiple EMR applications – Clinical Notes, Flow Sheets, GDMS. "Smart" paper questionnaires are implemented in Teleform™ software. Online forms are custom coded in java/.net. • Used in all practice settings – Hospital, Outpatient, Telephonic, at Home • Point of Care capability – printing of smart form, scanning of smart form • HL7 – Compliant data • Wireless capable (via Browser) • Flexible use – but flexibility requires training • Questionnaires are automatically scored and the results made immediately available for provider review. • Questionnaire results are integrated into the patient's EMR, patient responses may be reviewed and acknowledged, or reviewed and modified by provider. • Structured data elements and documentation are stored both clinically and made available via Enterprise Data Trust (EDT) to researchers. Multiple studies have used the data for validation and population based research. • Available in English,	• Demonstrated cost savings both in dollars and FTE • Data delivery is flexible – push, pull, stored procedure, web services, etc. Flexibility requires custom work. However some methods are very "reusable". Web services delivery used for GDMS, EDT. Push to G.E. Flow Sheets – usable for PHQ-9, ACT, Meaningful use. • 80%+ completion rates for PFH and CVI (adult/pediatric) – majority in paper. Patient Portal based online completion increasing ~15% in Rochester, ~50% in Florida. • 5-point stress scale validated against SCL-90 • Patient reported "Diagnosis" validated against Medical Record "Diagnosis"

			Spanish, and Arabic – answer recording structure allows conversion of answers to English for medical staff	
Northwestern Univ.	Personal Communication	No In process of implementing ePRO data collection with iPads in the clinic		
Johns Hopkins	Snyder et al[58]	No. Trying to find funding for phase 2 to assess website use (pt completion and clinician use), usefulness, acceptability. Plan to do at Hopkins with breast and prostate cancer pts.	• Prototype website to collect PROs in oncology clinic setting . • Microsoft ASP.Net and SQL database • Clinicians could assign questionnaires to pts and schedule frequency of completion. • System had built in calendar function to automatically generate emails to pts to alert them when time to complete the forms. • Pts could enter responses, submit comments, and view results to responses. Pts can see graphs of their responses over time and get explanations of their scores. • Clinicians received a text and graphical view of pt responses and scores over time. Responses made available in the EMR so they can view clinical and PRO information in one place	• Clinicians reported that the website could improve clinical practice if it was not burdensome. • Clinicians were most interested in tracking change over time. • Patients were interested in using the website because of the potential to facilitate communication with their clinicians. • Patients emphasized the importance of short and simple surveys and a user-friendly interface. • Usability testing suggested that patients had few problems accessing and using the site.
Duke Univ. & West Memphis Clinic	Abernethy et al[59]	No Duke Yes West Memphis Clinic	• Modified the PACE™ System (which stands for Patient Assessment Care and Education), developed by Supportive Oncology Services, Inc. (SOS, Inc. Memphis, TN). • The software comprises several clinical tools, including a Review of Systems data-collection tool, the Patient Care Monitor (PCM), and an electronic patient-education library, the Cancer Support Network. • The PCM software	• Patients responded favorably to use of etablets • They liked educational materials included on the tablet • Issues with dead zones • Interface issues with Duke IT and vendor- due to confidentiality concerns, could not allow vendor to use its on wireless system at Duke • Even though this PACE system had been embraced by community practices, the Duke oncologists were concerned that

			currently gathers patient-reported demographics, illness, symptoms, performance status, and QOL data. • Added additional QOL surveys into system.	implementation would add to their workload even though prior studies did not show this • Physicians were also concerned that they would have to assume the study nurse duty of reviewing thresholds, identifying pts with critical needs, and guiding patients in system use
Dana Farber Cancer Institute	Berry[60] Wolpin[61]	ePRO is being implemented in clinical practice at DFCI on a clinic-by-clinic basis; sarcoma clinic is the closest to full implementation).. There has not been a system wide decision to implement ESRA-C (my tested program) or any other ePRO approach as yet.	• Software developed at DFCI • Development time for ESRA-C was approximately 6 months and involved rapid prototyping and extensive testing following the usability engineering lifecycle proposed by Mayhew. • Usability testing was also conducted with a sample of proxy patients at a community center for adults with literacy needs, and minor revisions were made based on these results • The DHAIR platform was built on an open-source architecture comprising a Linux Operating System, Apache Web server, MySQL database system, and the PHP or PERL or Python programming languages (LAMP). An administrative interface provided a survey editing environment for researchers where questions and response options could be entered and immediately deployed. Options for layout, question branching, forced response, and user control were also available within this interface.	• Primary finding was that participants were able to use ESRA-C quickly and without difficulty in a real-world clinical setting and that they were quite satisfied with the ESRA-C platform. • Fact that nearly 20% answered questions out of sequence points to the need for designing flexible navigation systems, notably, providing mechanisms for returning to prior questions to re-evaluate responses. • Mean survey administration time of 15 minutes 20 seconds is feasible within a busy clinical setting.

Memorial Sloan Kettering	Basch[62] Vickers[63]	Yes- in postprostatectomy patients	• Web based STARS system • Interacts with EMR to identify surgery dates and pt email address to allow pts to be automatically linked to Web survey . • Clinicians can view numerical and graphical summaries of pt functioning over time from within EMR. • Pt can also see info thru the STAR system Web interface • Can see the avg functional improvement of pts with similar characteristics and what functioning will likely be in future. • Programmed to generate summaries and run prediction models using pt reports of functioning	
Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital	Aaronson[8] Personal communication	Currently setting up a patient portal (individualized patient homepage) whereby, among other things, patients will complete online PRO questionnaires. It will have other components consistent with a survivorship care plan. This will become part of routine care, although it is being developed with a grant from the Dutch Cancer Society		
United Kingdon-Scotland	McCann[64]	Not reported	• mobile phone-based advanced symptom management system (ASyMS)	• 97% of the patients felt that they had received enough training to use handset on own and felt either comfortable (19%) or very comfortable (81%) using the ASyMS© system. • Patients were extremely positive about the use of the ASyMS© handset to

				record their symptoms • 91% of patients felt that using the ASyMS© handset had helped in symptommanagement.
Cleveland Clinic	Gurland[65] Personal Communication	KP is rolled out in 56 centers in 9 of our Institutes Uses etablets Pilot initially done in surgical practice	• The Knowledge Program (TKP), a proprietary software developed at Cleveland Clinic • can be divided into 2 distinct categories: HSM and structured documentation. HSM is the compilation of validated questionnaires used to collect patient responses during an outpatient visit. • The questionnaires are automatically scored and the results made immediately available for provider review. • Questionnaire results are integrated into the patient's EHR on provider approval (Epic, Verona Wi). • The structured documentation component is focused on discrete storage of information for later retrieval and clinical outcomes research.	• that tablet technology, including the use of TKP software, can be successfully integrated into a colorectal surgery practice. • By identifying and educating the key personnel in the patient visit process, as well as requesting that patients arrive 30 minutes before their scheduled appointment time, we were able to minimize disruption to the clinical flow. • Although the personal health record (PHR) was not used for this pilot project, patients also have the opportunity to fill out questionnaires at home before their scheduled visit using the PHR, thus avoiding coming to the clinic earlier than their scheduled appointment. • The PHR can be used to send out questionnaires at regular intervals following surgery, or in cases where patients would not be scheduled for in-person follow-up. • By use of HL7 messaging and wireless technology based on scheduled visit type codes, the amount of human intervention involved in questionnaire

				distribution is reduced, thus decreasing the need for personnel and improving efficiency. • vastly superior return rate for electronic questionnaires (96%) compared with that of paper questionnaires (25%). We speculate that the dramatic increase in response rates was associated with the change in workflow and the expectation communicated to the patient that their questionnaires were part of their clinical care, unlike paper forms that were strictly for research purposes. • individual questionnaire responses varied from 95% to 20%. Sexual function questionnaires had the poorest return rate. This analysis was not performed for the paper forms.

Appendix C

PROMIS v.1.0 - GLOBAL

Global Items

Please respond to each item by marking one box per row.

		Excellent	Very good	Good	Fair	Poor
	In general, would you say your health is:.....	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
	In general, would you say your quality of life is:	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
	In general, how would you rate your physical health?.....	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
	In general, how would you rate your mental health, including your mood and your ability to think?.....	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
	In general, how would you rate your satisfaction with your social activities and relationships?.....	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
	In general, please rate how well you carry out your usual social activities and roles. (This includes activities at home, at work and in your community, and responsibilities as a parent, child, spouse, employee, friend, etc.).....	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
		Completely	**Mostly**	**Moderately**	**A little**	**Not at all**
	To what extent are you able to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair?.....	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1

In the past 7 days…

		Never	Rarely	Sometimes	Often	Always
Global10	How often have you been bothered by emotional problems such as feeling anxious, depressed or irritable?.....	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

		None	Mild	Moderate	Severe	Very severe
Global08	How would you rate your fatigue on average?....	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Global07	How would you rate your pain on average?.....	☐ 0 **No pain**	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10 **Worst imaginable pain**